other P domain residues. A model of the GII.10 capsid was built as described above, and the relative solvent accessibility (SASA) of the P domain residues in the context of the capsid was computed using the ASAView server (2). Residues were then divided into 10 categories according the estimated relative SASAs, as follows: >0.45, >0.40, >0.35, >0.30, >0.25, >0.20, >0.15, >0.10, >0.05, and all SASA. For each SASA category, the average of the already-computed average per-residue RMSDs was determined separately for the binding site residues and for the nonbinding site residues.

Protein structure accession numbers. Atomic coordinates and structure factors were deposited in the Protein Data Bank under the following IDs: for GII.10, 3ONU (unliganded), 3PA1 (A-trisaccharide), 3Q38 (B-trisaccharide), 3Q39 (H type 2-disaccharide), 3Q3A (H type 2-trisaccharide), 3ONY (Le^b-fucose) 3Q6Q (Le^a disordered), 3Q6R (Le^x disordered), and 3PA2 (Le^y-tetrasaccharide), and for GII.12, 3R6J (unliganded) and 3R6K (B-trisaccharide).

RESULTS

Unbound structure of the GII.10 P domain. The GII.10 P domain MBP fusion protein was expressed at a level of ~10 mg/liter in *E. coli*. The cleaved GII.10 P domain formed rectangular plates that diffracted to better than 1.5-Å resolution (Table 1). A molecular replacement solution with the previously determined GII.4 P domain (4) was obtained in space group $P2_1$, with one P domain dimer in the asymmetric unit (Fig. 1A and B). Refinement of the GII.10 structure led to an R_{work} value of 0.151 (R_{free} = 0.167) and well-defined density for most of the P domain dimer (Table 1). Following the nomenclature established by Prasad and colleagues (30), the GII.10 P1 subdomain was located between residues 222 to 277 and residues 427 to 549, whereas the P2 subdomain was between residues 278 and 426. The GII.10 P1 subdomain was formed primarily by a single α-helix, which was flanked by seven antiparallel β-strands (Fig. 1B). The GII.10 P2 subdomain contained 12 antiparallel β-strands, 6 from each subunit, which formed 2 antiparallel β-sheets (Fig. 1B). Overall, the secondary structure of the GII.10 P domains was highly reminiscent of previously published GI and GII structures (4, 30). On one of the asymmetric unit monomers, residues 344 to 351 (chain B) were disordered; these disordered residues were not modeled into the GII.10-apo structure.

Unbound structure of the GII.12 P domain. The GII.12 P domain MBP fusion protein was expressed at a level of ~2 mg/liter in *E. coli*. The cleaved GII.12 P domain formed rectangular parallelepipeds that diffracted to 1.75-Å resolution (Table 2). The GII.12 P domain structure was determined by molecular replacement with the GII.10 P domain; structure solution indicated that the space group was $C222_1$, with one P domain monomer in the asymmetric unit (Fig. 1C, with its monomeric P1 and P2 subdomain partners shown in green and cyan, respectively). Refinement of the GII.12 structure led to an R_{work} value of 0.185 (R_{free} = 0.203) and well-defined density for most of the P domain monomer (Table 2). The GII.12 P1 subdomain was located between residues 222 to 277 and residues 414 to 536, whereas the P2 subdomain was between residues 278 and 413.

Comparisons of unbound structures of the GII.10, GII.12, GI.1, and GII.4 P domains. Despite the great genetic diversity of noroviruses, the GII.4 strains have been responsible for the majority of outbreaks around the world over the past 10 or so years (25, 35, 36). To examine whether the rare versus outbreak status had bearing on the overall structures, we compared rare and outbreak GII strains. The P domains from rare GII.10 and GII.12 were highly similar in structure, with a root mean square deviation (RMSD) for Cα atoms of 0.64 Å. However, in addition to their shared rare status, they were also more closely genetically related to each other than to the GII.4 outbreak strain. Pairwise analysis of RMSD differences in the P domain structures (Fig. 1) found that the three GII P domain structures, two rare and one outbreak, were more similar to each other than to the GI structure. Overall structural differences thus appeared to reflect genetic distance (see Fig. S1A in the supplemental material) rather than rare or outbreak status.

Structures of HBGA H type 2-trisaccharide and -disaccharide bound to the GII.10 P domain. HBGAs are a group of short oligosaccharides that are expressed in a polymorphic manner on cell surfaces or found as free antigens and have been shown through a number of studies, including the aforementioned crystallographic ones, to interact with norovirus (Fig. 2) (11, 19). HBGAs are generated from a number of different precursor disaccharides, with additional saccharides added by enzymes, which are variably present in the human population (see Fig. S2 in the supplemental material) (22). One distinction is made by the presence of α1,2fucosyltransferase, which adds a terminal αfucose1-2 unit; HBGAs with this saccharide are termed secretors, while those missing the terminal αfucose1-2 are termed nonsecretors.

Because the GII.10 P domain protein was expressed to larger amounts and crystals diffracted to higher resolution than those of GII.12, we chose to examine first the GII.10 P domain by X-ray crystallography in complex with a panel of HBGAs (see Fig. S2 in the supplemental material) representing an assortment of secretor and nonsecretor HBGAs. The secretor HBGAs used were H type 2-disaccharide, H type 2-trisaccharide, A-trisaccharide, B-trisaccharide, Le^y-tetrasaccharide, and Le^b-tetrasaccharide, whereas the nonsecretor HBGAs used were Le^a-trisaccharide and Le^x-trisaccharide.

The HBGA H type 2-trisaccharide is α-L-fucose-(1-2)-β-D-galactose-(1-4)-2-*N*-acetyl-β-D-glucosamine, which is the first secretor in one of the major biosynthetic HBGA pathways (see Fig. S2 in the supplemental material). Cocrystallization of the GII.10 P domain with H type 2 resulted in $P2_1$ crystals that diffracted to 1.40 Å, with cell constants virtually isomorphous with those of the unbound crystals (Table 1). Structure solution and refinement with the unbound P domain resulted in a single clearly defined patch of electron density that spanned two P domain monomers (Fig. 2A and 3A). Placement of the trisaccharide was assisted by a well-defined fucose density, which led to an unambiguous orientation of this HBGA. Refinement led to an R_{work} value of 0.169 (R_{free} = 0.188) and well-defined density for all of the saccharide units (Fig. 3A). No unassigned electron density was observed in the corresponding position of the HBGA on the P domain dimer, around the molecular 2-fold. Inspection of the lattice indicated a lattice contact at this position, which would occlude the presence of a second HBGA molecule (see Fig. S3A in the supplemental material).

The fucose showed the most well-defined density and was fixed by a network of P2 subdomain hydrogen bonds, two contributed by the side chain of Asp385, two by the side chain of Arg356, and one by the main chain of Asn355 (Fig. 3B; see also Fig. S1B in the supplemental material). A sixth hydrogen bond was contributed from the backbone of Gly451 from

TABLE 1. Data collection and refinement statistics for structures of the GII.10 norovirus P domain alone and with various HBGAs

Statistics	Value(s) for[a]:								
	026 Le[b] (monoglycan) 3ONY	026 apo (no glycan) 3ONU	026 A (triglycan) 3PA1	026 B (triglycan) 3Q38	026 H type 2 di (diglycan) 3Q39	026 H type 2 tri (triglycan) 3Q3A	026 Le[a] (disordered) 3Q6Q	026 Le[x] (disordered) 3Q6R	026 Le[y] (tetraglycan) 3PA2
Data collection									
Space group	$C222_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions									
a, *b*, *c* (Å)	80.53, 115.91, 267.80	65.22, 79.11, 69.30	65.19, 78.99, 70.10	65.23, 79.03, 69.64	65.12, 78.77, 69.49	65.17, 79.10, 68.91	65.15, 78.96, 69.24	65.28, 79.02, 69.70	65.27, 79.13, 69.07
α, β, γ (°)	90, 90, 90	90, 99.65, 90	90, 101.06, 90	90, 99.84, 90	90, 99.93, 90	90, 99.54, 90	90, 99.67, 90	90, 99.87, 90	90, 100.27, 90
Resolution range (Å)	50–1.85 (1.92–1.85)	50–1.40 (1.45–1.40)	30–1.48 (1.52–1.48)	30–1.28 (1.31–1.28)	30–1.25 (1.28–1.25)	30–1.40 (1.44–1.40)	30–1.43 (1.47–1.43)	30–1.40 (1.44–1.40)	30–1.48 (1.52–1.48)
R_{sym}	8.4 (58.9)	7.0 (44.9)	6.6 (65.7)	4.1 (72.5)	5.2 (67.4)	4.0 (67.2)	3.9 (65.3)	4.4 (71.0)	5.7 (64.7)
$I/\sigma I$	22.8 (2.2)	18.4 (2.0)	12.5 (2.3)	18.9 (2.1)	16.1 (2.3)	21.3 (2.3)	22.0 (2.2)	22.2 (2.4)	16.3 (2.4)
Completeness (%)	99.1 (93.3)	95.2 (69.6)	99.8 (99.9)	99.8 (99.8)	99.7 (99.7)	99.8 (99.9)	96.6 (94.1)	99.6 (99.7)	99.7 (99.8)
Redundancy	6.8 (5.2)	3.6 (2.7)	3.8 (3.7)	3.7 (3.6)	3.7 (3.4)	3.7 (3.7)	3.8 (3.5)	3.7 (3.7)	3.7 (3.7)
Refinement									
Resolution range (Å)	30–1.85	30–1.40	30–1.48	30–1.28	30–1.25	30–1.40	30–1.43	30–1.40	30–1.48
No. of reflections	99,983	122,330	115,862	178,392	190,020	135,210	123,157	136,413	114,698
R_{work}/R_{free}	0.164/0.189	0.151/0.167	0.178/0.198	0.167/0.181	0.168/0.182	0.169/0.188	0.177/0.188	0.165/0.178	0.185/0.204
No. of atoms	7,949	5,755	5,814	5,883	5,798	5,779	5,721	5,893	5,778
Protein	7,149	4,814	4,927	4,915	4,861	4,882	4,876	4,826	4,870
Ligand/ion	73	32	110	143	134	149	88	100	143
Water	727	909	777	825	803	748	757	967	765
B-factors									
Protein	40.4	20.9	20.8	19.3	17.5	21.4	19.9	18.9	19.6
Ligand/ion	41.1	21.5	28.1	30.6	29.5	28.7	27.6	24.7	23.7
Water	39.5	32.9	33.2	32.4	30.6	33.9	33.4	35.1	32.2
RMSD									
Bond length (Å)	0.007	0.011	0.008	0.011	0.018	0.009	0.005	0.006	0.007
Bond angle (°)	1.033	1.331	1.155	1.357	1.758	1.244	0.984	1.062	1.120

[a] Each data set was collected from a single crystal. Values in parentheses are for the highest-resolution shell.

TABLE 2. Data collection and refinement statistics for structures of the GII.12 norovirus P domain alone and with triglycan HBGA type B

Statistics	Value(s) for[a]: Hiro apo (no glycan) 3R6J	Hiro B (triglycan) 3R6K
Data collection		
Space group	$C222_1$	$C222_1$
Cell dimensions		
a, *b*, *c* (Å)	73.01, 99.20, 77.60	73.39, 100.28, 82.15
α, β, γ (°)	90, 90, 90	90, 99.84, 90
Resolution (Å)	50–1.85 (1.81–1.75)	50–1.60 (1.66–1.60)
R_{sym}	5.0 (44.9)	8.6 (39.5)
$I/\sigma(I)$	28.0 (2.5)	18.3 (2.0)
Completeness (%)	98.9 (92.0)	91.7 (61.2)
Redundancy	5.1 (3.9)	5.9 (3.1)
Refinement		
Resolution (Å)	25–1.75	25–1.60
No. of reflections	28,478	36,791
R_{work}/R_{free}	0.185/0.203	0.219/0.237
No. of atoms	2,487	2,607
Protein	2,359	2,338
Ligand/ion	4	41
Water	124	228
B-factors		
Protein	45.8	32.1
Ligand/ion	50.7	66.9
Water	42.0	33.8
RMSD		
Bond length (Å)	0.004	0.005
Bond angle (°)	0.887	0.930

[a] Each data set was collected from a single crystal. Values in parentheses are for the highest-resolution shell.

across the P domain dimer interface, with the aromatic ring of Tyr452 packing over the fucose methyl. Both Gly451 and Tyr452 are located on a loop that extends from the P1 subdomain to form part of the P domain dimer interface (Fig. 1E). Meanwhile, the galactose was fixed by one hydrogen bond, and the *N*-acetyl-glucosamine by three, contributed by a mix of backbone and side chain interactions, including Lys449 on the aforementioned P1-interface loop (Fig. 3B; see also Fig. S1B).

To better understand H type 2 recognition, we also determined the structure of an H type 2-disaccharide [α-L-fucose-(1-2)-β-D-galactose] in complex with the GII.10 P domain (Table 1). The fucose appeared well ordered, but the galactose ring was substantially less well defined (Fig. 3C). Apparently the single observed hydrogen bond to the galactose ring in the trisaccharide structure was not sufficient to fix the galactose in the disaccharide structure when not also sandwiched by an *N*-acetylglucosamine, as in the H type 2-trisaccharide (Fig. 3).

Overall, the unbound and H type 2-bound structures of the GII.10 P domain were virtually indistinguishable, except that in the bound structures, saccharides replace a number of surface waters. Within the bound H type 2 HBGAs, the primary interactions were observed to be through the terminal αfucose1-2 moiety, which was tightly held by both hydrophobic and hydrophilic interactions at the P domain dimer interface and involved the P1-interface loop from one monomer and the P2 subdomain from another monomer (Fig. 2A and 3).

Structure of HBGA Le^y-tetrasaccharide bound to the GII.10 P domain. The Le^y-tetrasaccharide HBGA is α-L-fucose-(1-2)-β-D-galactose-(1-4)-2-*N*-acetyl-β-D-glucosamine-(3-1)-α-L-fucose, which is the product of α1-3fucosyltransferase on H type 2-trisaccharide HBGA (see Fig. S2 in the supplemental material). Cocrystallization of the GII.10 P domain with Le^y resulted in $P2_1$ crystals that diffracted to 1.48 Å, with cell constants virtually isomorphous with those of the unbound and H type 2-bound crystals (Table 1). Similar to the H type 2 structure described above, the Le^y complex structure solution and refinement resulted in a single patch of electron density, which overlapped with the position of the αfucose1-2 in the H type 2 complex structure (Fig. 2A and 4A). The Le^y-tetrasaccharide was tested in the following two orientations: either with αfucose1-2 or with αfucose1-3 placed in the P domain interface. Only the αfucose1-2 placement refined well. Refinement led to an R_{work} value of 0.185 (R_{free} = 0.204) and well-defined density for all of the saccharide units (Fig. 4A).

As described for the H type 2 complex structures, the αfucose1-2 of Le^y was fixed by a network of six hydrogen bonds, i.e., two by Asp385, two by Arg356, one by Asn355, and one by Gly451, and a Tyr452-hydrophobic interaction, (Fig. 4B; see also Fig. S1B in the supplemental material). The galactose of Le^y was fixed by one water-mediated hydrogen bond, the *N*-acetylglucosamine by two backbone hydrogen bonds, and the terminal αfucose1-3 by a hydrogen bond to the side chain of Trp381. Interestingly, the positions of the saccharides, other than αfucose1-2, in Le^y were quite different from those in H type 2 (Fig. 5A). In Le^y, the galactose kinks up away from the protein, the *N*-acetylglucosamine swivels closer to the protein, and the terminal αfucose1-3 ends up being positioned close to the location of the third saccharide (*N*-acetylglucosamine) from H type 2.

HBGA Le^b-tetrasaccharide bound to the GII.10 P domain as a single ordered fucose. The Le^b-tetrasaccharide HBGA is α-L-fucose-(1-2)-β-D-galactose-(1-3)-2-*N*-acetyl-β-D-glucosamine-(4-1)-α-L-fucose, which is the product of α1-4fucosyltransferase on H type 1-trisaccharide HBGA (see Fig. S2 in the supplemental material). Cocrystallization of the GII.10 P domain with Le^y resulted in C-centered orthorhombic crystals that diffracted to 1.85 Å, and structure solution with the unbound GII.10 P domain structure revealed the crystals to be in space group $C222_1$, with three monomers of the P domain in the asymmetric unit (see Fig. S3B in the supplemental material). These three monomers formed the previously observed dimer, with the monomer arranged around a crystallographic 2-fold, so that it also formed the standard dimer.

Refinement to an R_{work} value of 0.164 (R_{free} = 0.189) revealed that the molecular dimer and the crystallographic dimer were virtually identical to each other (RMSD = 0.20 Å) and to the unbound dimer (RMSDs of 0.19 and 0.21 Å for the molecular and crystallographic dimer, respectively). Each of the three independent monomers contained a single somewhat poorly ordered αfucose1-2 (average B value of 49 Å^2), held in place by the standard six hydrogen bonds that spanned between two P domain monomers (Fig. 2A). Notably, other than this single fucose, no additional saccharides were observed (Fig. 4C and D).

Comparison of the structures of the H type 2-di- and -trisaccharide HBGAs indicated that without a third saccharide,

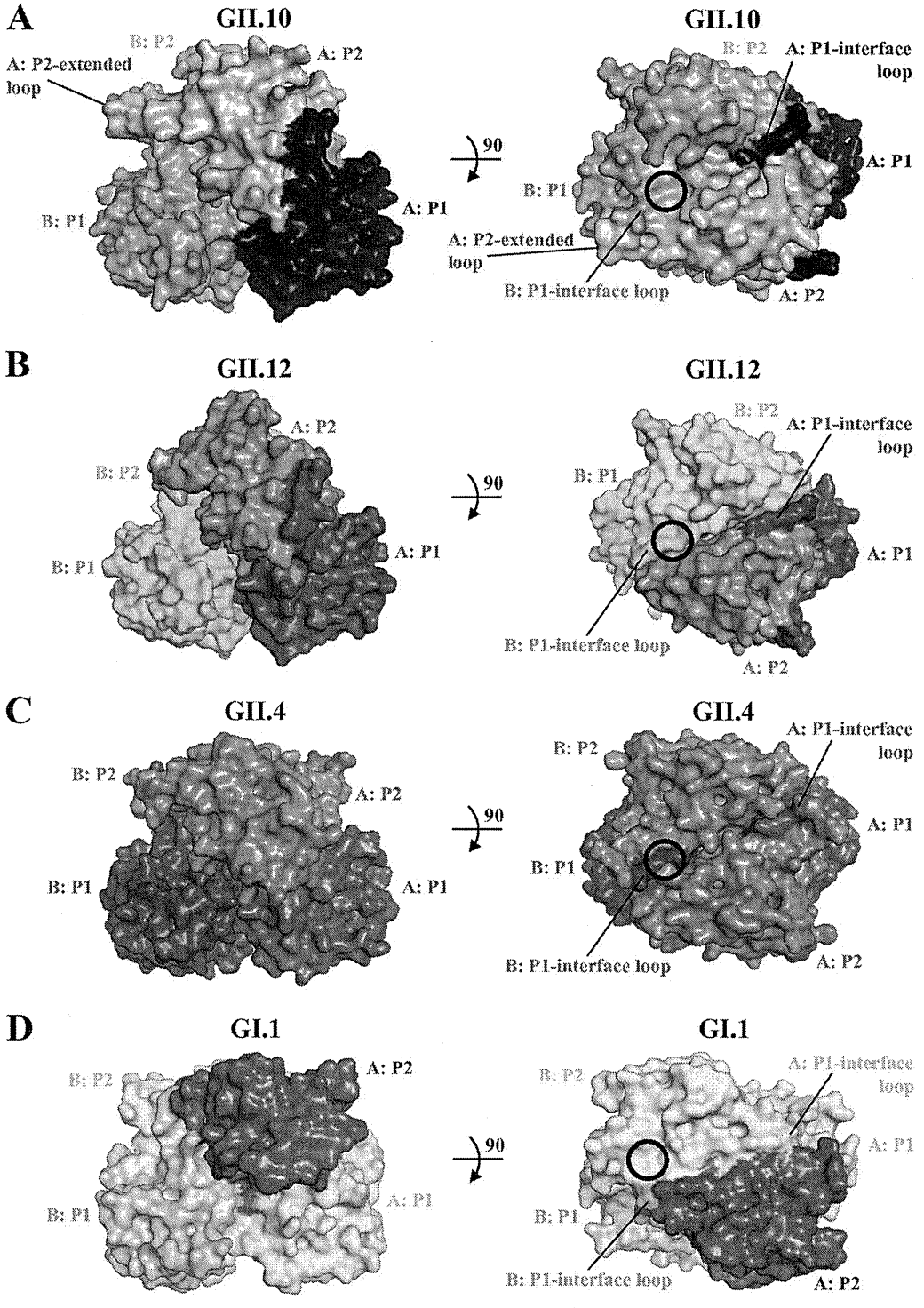

FIG. 2. Surface comparisons of the GII.10 (PDB ID, 3ONU), GII.12, GII.4 (PDB ID, 2OBR), and GII.1 (PDB ID, 2ZL5) P domain dimer structures. The GII HBGA binding sites (black circles in panels A to C) involve a dimeric capsid interface that is formed primarily by the P2 subdomain and includes a P1-interface loop, whereas the GI HBGA binding site (black circle in panel D) is monomeric, involves only a single P2 subdomain, and makes no contact with the P1 subdomain. (A) The GII.10 P2 subdomain had an amino acid insertion (relative to those of the other GII sequences), which corresponded to a P2-extended loop. (B) The GII.12 P2 subdomain was somewhat unlike the other two GII surfaces, having a more pointed P2 subdomain. (C) The GII.4 P2 subdomain was more similar to that of GII.10 but had a less pointed P2 subdomain top surface. (D) The GI.1 P domain appears somewhat flatter than that of the GII structures.

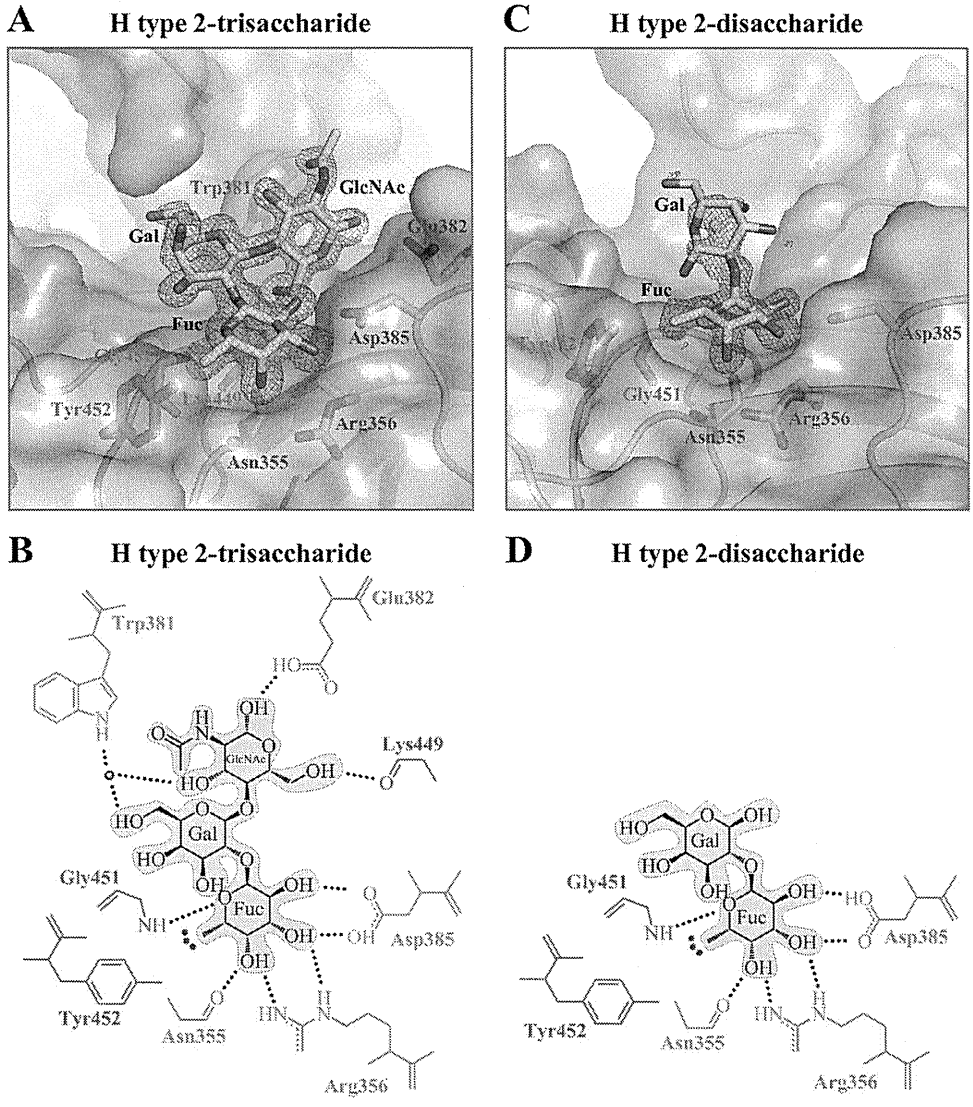

FIG. 3. GII.10 P domain and H type 2 (trisaccharide and disaccharide) interactions. The H type 2-tri- and -disaccharide binding site is at the same location on the P domain and utilizes identical residues to bind the terminal αfucose1-2 saccharide. (A) Close-up surface and ribbon representation of the GII.10 P domain (colored as described in the legend to Fig. 1B) showing the bound H type 2-trisaccharide (cyan) and electron density map contoured at 1.0 sigma. (B) GII.10 P domain and H type 2-trisaccharide hydrophilic and hydrophobic interactions (colored as described in the legend to Fig. 1B). The HBGA outline was shaded in blue, the black dotted lines represent the hydrogen bonds, the red dotted line represents the hydrophobic interaction from Tyr452, and the sphere represents water molecules. For simplicity, only the backbone was shown for residues that were backbone mediated. Hydrogen bond distances were less than 3.2 Å, though the majority was ~2.8 Å. (C) Close-up surface and ribbon representation of GII.10 showing the bound H type 2 disaccharide (cyan) and the electron density map at 1.0 sigma. (D) GII.10 P domain and H type 2-disaccharide hydrophilic and hydrophobic interactions.

the intervening galactose became partially disordered (compare Gal in Fig. 3A and C). Moreover, examination of the differences between the Le^b and Le^y chemistries indicated that the differences of these two could be envisioned as a swapping of the chemistries around the critical third saccharide ring, such that the two hydrogen bonds which are made at the first and second positions of that ring in the well-ordered Le^y-bound HGBA would be disrupted (compare GlcNAc in Fig. 4B and D). Thus, while we could not rule out completely different potential orientations for the bound Le^y HBGA, analysis of the other bound HBGAs indicated that only the αfucose1-2 of Le^b could bind in a manner similar to that of Le^y, consistent with the singly ordered fucose that was observed.

Structures of HBGA type A- and B-trisaccharides bound to the GII.10 P domain. The type A-trisaccharide HBGA is α-L-fucose-(1-2)-β-D-galactose-(3-1)-2-*N*-acetyl-α-D-galactosamine, whereas the type B-trisaccharide HBGA is the same as type A, except for a terminal α-D-galactose instead of an *N*-acetylgalactosamine [α-L-fucose-(1-2)-β-D-galactose-(3-1)-α-D-galactose]. Both of these HBGAs have the H type 2-disaccharide as a precursor (see Fig. S2 in the supplemental material). Cocrystallization of the GII.10 P domain with

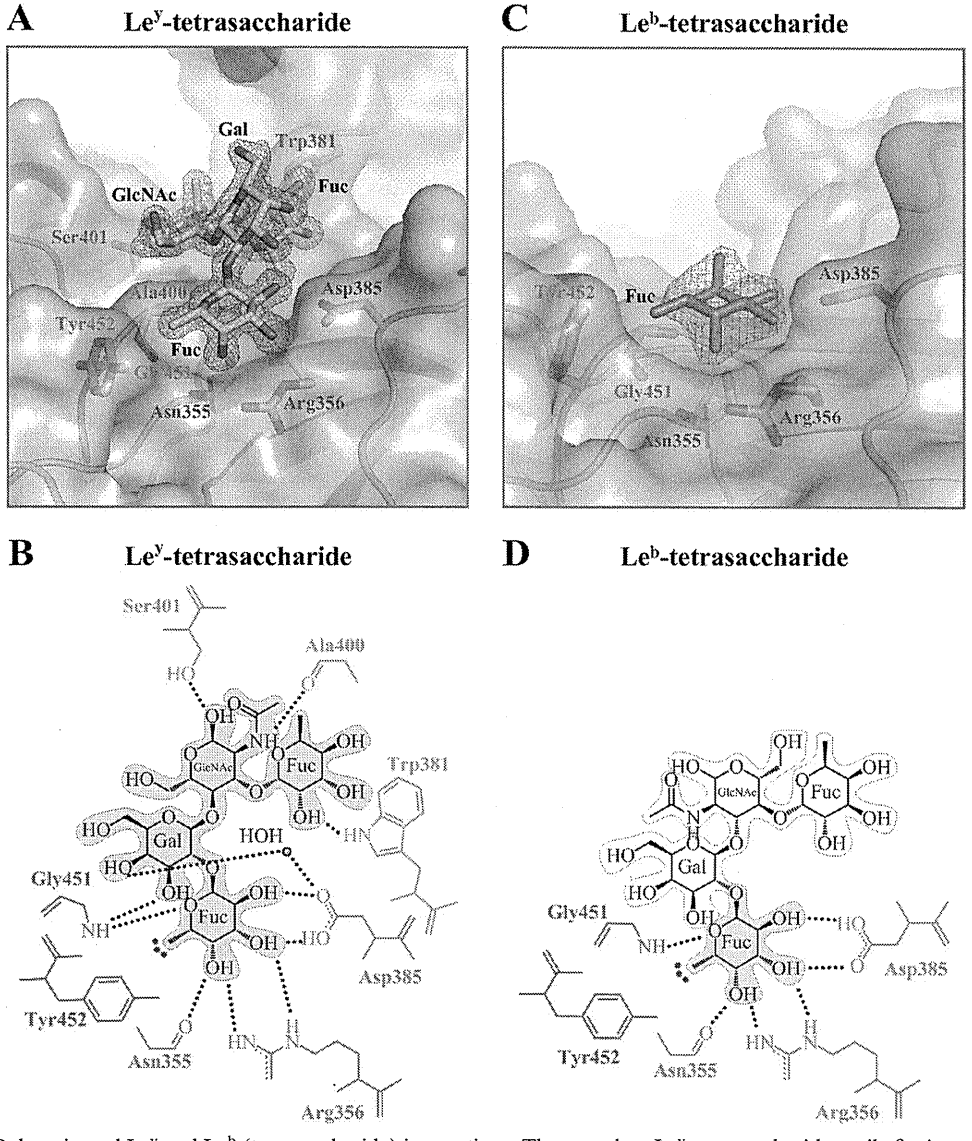

FIG. 4. GII.10 P domain and Le^y and Le^b (tetrasaccharide) interactions. The complete Le^y-tetrasaccharide easily fits into electron density and shows extensive hydrogen bonding interactions, whereas only αfucose1-2 of Le^b can be fit into the observed electron density; these differences in bound HBGA structure are likely the consequences of different glycosidic bonds on the third saccharide ring (see the text). (A) Close-up surface and ribbon representation of the GII.10 P domain showing the bound Le^y-tetrasaccharide (green) and the electron density map contoured at 1.0 sigma. (B) GII.10 P domain and Le^y hydrophilic and hydrophobic interactions (colored as described in the legend to Fig. 1B). (C) Close-up surface and ribbon representation of the GII.10 P domain showing the bound Le^b-tetrasaccharide (blue) and the electron density map at 1.0 sigma. (D) GII.10 P domain and Le^b-tetrasaccharide hydrophilic and hydrophobic interactions. The HBGA subunits that could not be fitted were outlined in light blue.

types A and B also resulted in $P2_1$ crystals that diffracted to 1.48 and 1.28 Å, respectively (Table 1). Similar to the structures described above, type A and B complex structure solutions and refinements resulted in a single patch of electron density, which overlapped with the position of the αfucose1-2 in the H type 2 complex structures (Fig. 2A). Placement of the αfucose1-2 of types A and B at the P domain interface allowed for the other two saccharides to be easily built into the remaining density. Refinement led to R_{work} values of 0.178 and 0.167 (R_{free} = 0.198 and 0.181) for type A and B bound structures, respectively, and well-defined density for all of the saccharide units (Fig. 6).

In addition to the six hydrogen bonds described above, αfucose1-2 was fixed by another water-mediated hydrogen bond to Lys449 (Fig. 6B and D). In total, five hydrogen bonds were contributed by one monomer of the P2 subdomain (Asn355, Arg356, and Asp385), and two were contributed by the P1-interface loop on the other monomer (Lys449 and Gly451), which also contributed the Tyr452-hydrophobic interaction (see Fig. S1B in the supplemental material). For type A, the galactose was fixed by one backbone-mediated hydrogen bond to Gly451, and the *N*-acetylgalactosamine by two water-mediated hydrogen bonds to Glu382. For type B, interactions were virtually identical,

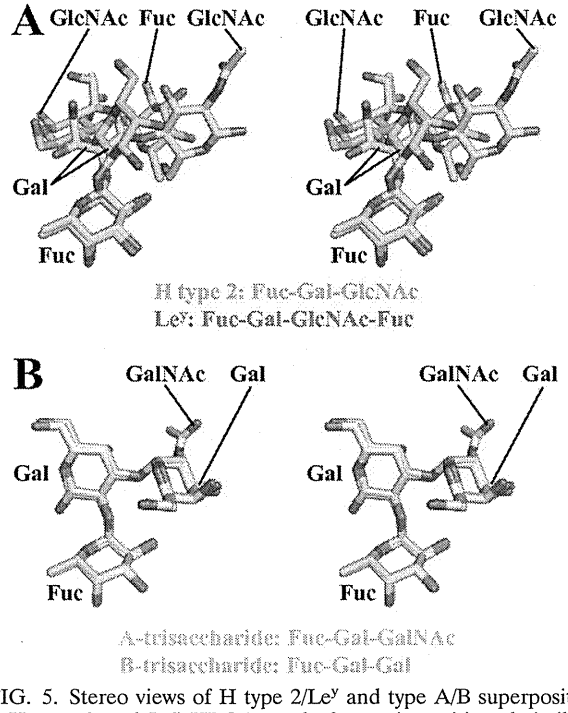

FIG. 5. Stereo views of H type 2/Le^y and type A/B superposition. For H type 2 and Le^y HBGAs, only fucose is positioned similarly, whereas for type A and B HBGAs, all saccharides are held in practically identical positions. (A) Stereo view of the H type 2 (cyan) and Le^y (green), showing the similar orientation of αfucose1-2 but the different positions of the other saccharides. (B) Stereo view of types A and B (yellow and pink, respectively), showing the similar orientations of each saccharide.

with the α-D-galactose also fixed by two water-mediated hydrogen bonds to Glu382.

In contrast to H type 2 and Le^y, types A and B bound in remarkably similar manners, with all atoms of fucose and galactose superimposing after alignment of P domain, with an RMSD of less than 0.01 Å (Fig. 5B).

Nonsecretor HBGAs Le^a- and Le^x-trisaccharides were not observed to bind to the GII.10 P domain. The HBGAs Le^a-trisaccharide and Le^x-trisaccharide are the product of the α1,3/4fucosyltransferase, which adds a terminal αfucose1-3/4 unit to the standard galactose-*N*-acetylglucosamine precursor. These HBGAs are termed nonsecretors because they lack a αfucose1-2 unit. Cocrystallization of these with the GII.10 P domain resulted in monoclinic crystals that diffracted to 1.40 and 1.43 Å for Le^a and Le^x, respectively, and molecular replacement and refinement revealed the standard $P2_1$ structure (Table 1), though in both cases, the patch of electron density was quite weak and no saccharide could be fitted (structures deposited without HBGA).

Structure of HBGA type B-trisaccharide bound to the GII.12 P domain. Having determined structures of the GII.10 P domain with a panel of HBGAs, we next turned to the GII.12 P domain. Cocrystallization of the GII.12 P domain with the type B-trisaccharide HBGA resulted in $C222_1$ crystals that diffracted to 1.60 Å (Table 2). Structure solution and refinement with the unbound GII.12 P domain resulted in a small patch of electron density, located at the P domain interface (Fig. 2B). Refinement led to an R_{work} value of 0.219 (R_{free} = 0.237). The fucose appeared very well ordered, while the two other saccharides were less well defined (Fig. 7A). The fucose was held in place by the standard six hydrogen bonds that spanned between two P domain monomers (Fig. 7B). However, in the case of GII.12, a main-chain hydrogen bond from cysteine (Cys345) replaced the GII.10 main-chain hydrogen bond from asparagine (Asn355).

Conservation of the HBGA binding motif in GII noroviruses. The structure of the outbreak GII.4 (VA387) strain of norovirus previously determined with HBGA type A- and B-trisaccharides closely resembles the GII.10 and GII.12 norovirus structures with HBGAs described here. Taken together, they reveal a coherent picture of HBGA recognition, dominated by αfucose1-2 binding, as observed by Tan et al. (41).

Of the 13 potential hydrogen bonds made by a terminal fucose, 6 are made by all 3 GII P domains in all 9 different HBGA P domain structures. These six, which are located in almost exactly the same places in all HBGA-bound structures, consist of five from a P2 subdomain and one from the P1-interface loop on another P domain monomer (Fig. 2A to C). These extensive contacts are quite specific for αfucose1-2, with αfucose1-3 unable to fit. The GII.10 and GII.4 interactions are further strengthened by a hydrophobic contact with the side chains of Tyr452 and Tyr443 on the P1-interface loop, respectively. Saccharides other than αfucose1-2 are attached in diverse ways, held in place by a rotating cast of surface residues.

To identify regions of high/low structural conservation, the six structures of GII.10 bound to different HBGA were further analyzed. Per-residue nonhydrogen atoms RMSDs were computed for each pair of structures, and the average RMSD among all structure pairs for each residue was obtained. The RMSD values for the GII.10 binding site residues were then compared to the RMSD values of nonbinding site residues, with a range of solvent accessibility cutoffs. In all cases, residues interacting with the different HBGAs were more conserved structurally as opposed to nonbinding site residues, though the average RMSD values were generally low for both sets of residues (see Fig. S4 in the supplemental material).

Sequence conservation of GII noroviruses and comparison with GI noroviruses. The conserved GII recognition of HBGAs requires conservation of interacting residues. To understand the effect on sequence conservation engendered by this conserved recognition, we aligned a panel of GII norovirus sequences onto the atomic-level structures of GII.10 norovirus and analyzed conservation of surface residues relative to HBGA recognition. The residues on the surface of the P domain corresponding to the outer surface of the capsid were substantially less conserved than the inward facing surface residues (Fig. 8A). On the outer facing surface, two major regions of high conservation were observed. These overlapped with the two dimer-equivalent regions that interact with αfucose1-2 of the HBGA (Fig. 8A, middle, and B). Notably, the residues forming the surface of the P domain that interacts with the peripheral saccharides were generally less conserved than the αfucose1-2-interacting residues (see Fig. S5 in the supplemental material). Thus, the structure-function relationships involved in HBGA recognition appear to be reflected in the conservation of the GII norovirus surface residues.

To test whether this conservation was indeed a reflection of

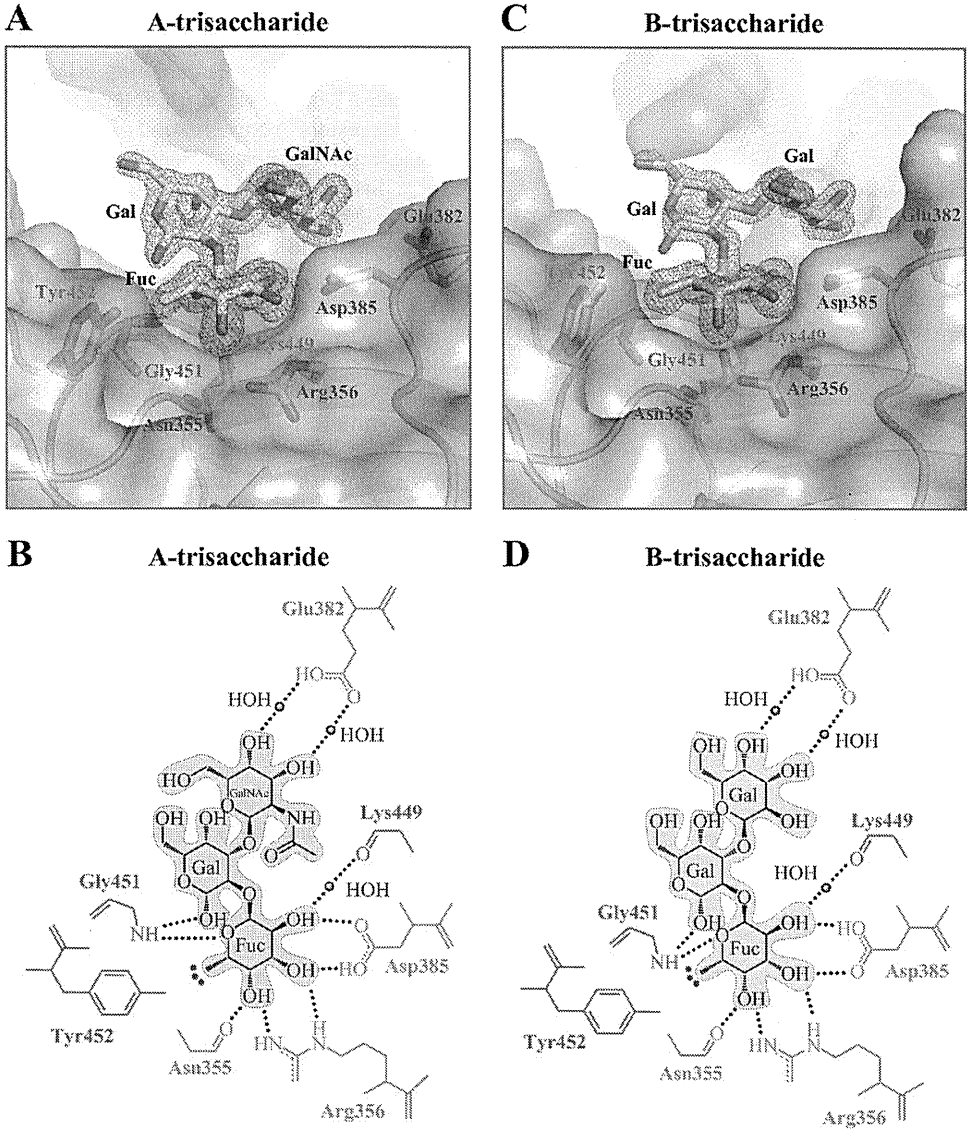

FIG. 6. GII.10 P domain and type A and B (trisaccharide) interactions. The GII.10 P domain interacts with type A and B HBGAs in virtually identical ways. (A) Close-up surface and ribbon representation of the GII.10 P domain showing the bound A-trisaccharide (yellow) and the electron density map contoured at 1.0 sigma. (B) GII.10 P domain and A-trisaccharide hydrophilic and hydrophobic interactions. (C) Close-up surface and ribbon representation of the GII.10 P domain showing the bound B-trisaccharide (pink) and the electron density map at 1.0 sigma. (D) GII.10 P domain and B-trisaccharide hydrophilic and hydrophobic interactions.

HBGA recognition, we aligned a panel of GI norovirus sequences (10) onto the previously determined structures (6) of GI.1 norovirus in complex with the HBGA type A and type H saccharides. The residues forming the surface of the GI P domain corresponding to the outer surface of the capsid were also substantially less conserved than the inward facing surface residues (see Fig. S6 in the supplemental material). On the outer facing surface, two regions of high conservation were observed. These overlapped with the dimer-equivalent regions on each monomer that interact with the HBGAs (Fig. S6). Notably, the surface patch formed by conserved residues in the GI noroviruses was in a different location than the patch in the GII noroviruses. In both cases, the sites of sequence conservation related to the regions involved in HBGA recognition, which is in agreement with previous observations (4, 6, 41). Thus, the structure-function relationships involved in HBGA recognition appear to be reflected in surface-residue conservation for both GI and GII noroviruses.

The region of high conservation on the GII.10 outer facing surface included an additional residue, His358, which was not part of the identified HBGA binding sites (see Fig. S7 in the supplemental material). In our structures and in the GII.4 structures determined previously (4), this residue was observed to make a potential hydrogen bond with the side chain of Asp385. The conservation of both Asp385 and His358 suggests that these two residues form a hydrogen bonding network that may be essential for HBGA binding of GII viruses. Due to its solvent exposure and adjacency to the

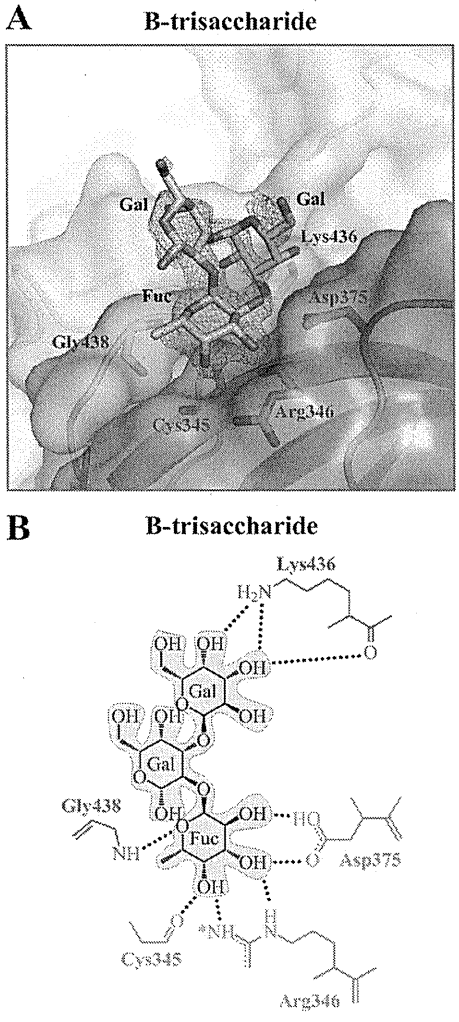

FIG. 7. GII.12 P domain and B-trisaccharide interaction. The GII.12 P domain binds αfucose1-2 of type B HBGA with hydrogen bonds similar to those of GII.10, except that the carbonyl of Cys345 replaces that of Asn355. (A) Close-up surface and ribbon representation of the GII.12 P domain (colored as described in the legend to Fig. 1C and shown as a dimer) showing the bound B-trisaccharide (pink) and the electron density map contoured at 1.0 sigma. (B) GII.12 P domain and B-trisaccharide hydrophilic interactions. The asterisk on Arg346 indicates that a hydrogen bond interaction was slightly longer (3.3 Å) than the other bonds, usually less than 3.1 Å.

fucose-binding site residues, it may be possible for His358 to also participate in direct binding interactions with some HBGAs. Likewise for GII.12, His348 (GII.12 numbering) was observed to form a similar hydrogen bond with the side chain of Asp375 (data not shown).

DISCUSSION

Viruses often use genetic variability to escape host recognition. Such variation, however, is limited by function: the virus cannot alter functionally critical elements while retaining infectivity. In particular, recognition of host factors, such as receptors or cofactors, generally requires regions on the outer surface of the virus to remain conserved. In the case of HIV-1, interaction with the CD4 receptor requires part of the HIV-1 gp120 envelope glycoprotein to remain conserved, and this same site is recognized by antibody VRC01, which is able to neutralize over 90% of circulating HIV-1 isolates (45, 48). In the case of influenza virus, interaction with the sialic acid receptor results in conservation of a small surface patch on the hemagglutinin trimer, and small molecules and antibodies that target this patch have been less successful at broadly neutralizing diverse strains of influenza virus (43, 44). With noroviruses, functional requirements related to HBGA recognition could potentially require substantial portions of the capsid surface to remain conserved and thereby serve as sites of vulnerability to small molecule- or antibody-mediated neutralization.

One way that noroviruses might alter such conservation requirements is by varying their modes of interactions with HBGAs. If different noroviruses were to use different modes of interactions, then different conservation schemes—and enhanced variation—would result. Indeed, different modes of HBGA are observed between the GI and GII genotypes of human noroviruses (3, 4, 6). The crystal structures obtained here from rare GII isolates (GII.10 and GII.12), however, show means of HBGA recognition virtually identical to those of the previously determined outbreak GII.4 structures (4). These results suggest that within GII, a single mode of recognition occurs.

The size of a HBGA is roughly half the size of an antibody epitope. If HBGA recognition were to require a conserved surface of roughly this size, such conservation could lead to significant vulnerability to antibody-mediated neutralization. Structure-function analysis of the GII.10 norovirus with a panel of HBGAs, however, indicates conserved binding at only one saccharide unit, terminal αfucose1-2, with variable recognition at peripheral saccharide units. Apparently, norovirus uses variation in human HBGAs, along with flexibility between saccharide units within each HBGA and variation in amino acid side-chain stereochemistry, so that the same amino acids can recognize diverse HBGAs in different ways. This allows the GII noroviruses to reduce the size of the conserved interaction surface to residues under a single critical saccharide rather than the entire HBGA. Nevertheless, this conserved surface defines a potential site of vulnerability on GII viruses (Fig. 8C) and may thus present a useful target for therapeutic and/or vaccine design efforts.

The HBGAs analyzed here represent only a fraction of known HBGAs (22). Those described here are involved in a primary major biosynthetic pathway, happen to be commercially available, and were described in a number of previous papers characterizing norovirus HBGA interactions (11–13, 19, 31, 37, 38). We provide definition for this panel with GII.10 and GII.12 noroviruses, with crystal structures at ~1.5-Å resolution. The high resolution revealed unexpected details. In the HBGA with H type 2-trisaccharide, the αfucose1-2 refined as an βfucose, nuclear magnetic resonance (NMR) analysis of the commercially obtained trisaccharide shows a mixture of at least four components, including a βfucose-containing impurity (data not shown). The impurities in the commercially available HBGAs may also explain some of the inconsistencies among the different laboratories, as recently reported (39). Nonetheless, as the αfucose-(1-2)-β-D-galactose disaccharide unit is common to most of the HBGAs analyzed here, the

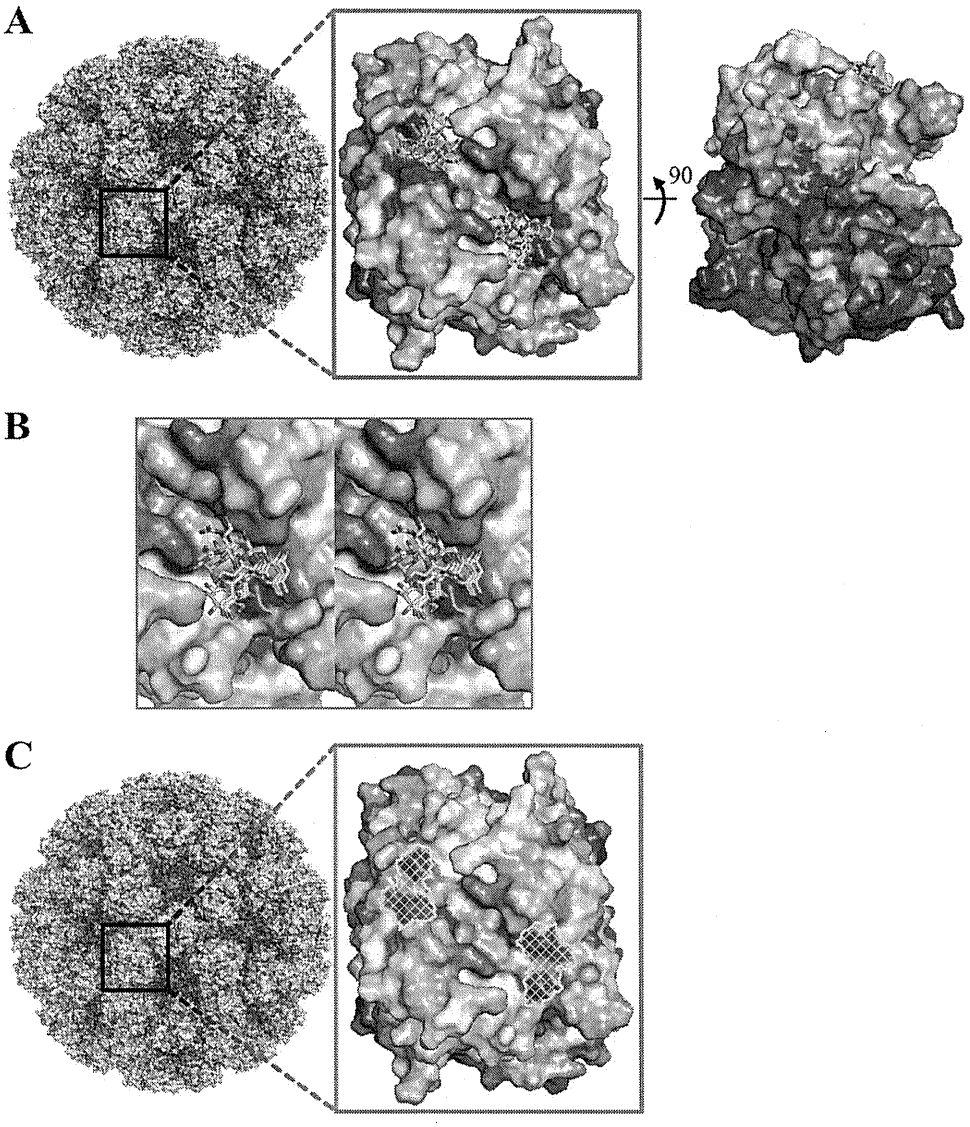

FIG. 8. Surface representations of GII amino acid conservation and putative site of vulnerability for GII noroviruses. Antigenic diversity of noroviruses is seen primarily on the outermost surface of the capsid, although patches of conservation on the top surface are observed. The most prominent of these patches correspond to the P domain-binding sites of the HBGAs described here. (A) An alignment of GII genotypes was used to map the amino acid conservation and variability on the GII.10 P domain dimer structure. The color-coded conservation ranged from a deep purple, represented by highly conserved amino acids, to white, represented by highly variable amino acids. GII conservation was mapped onto a model of the viral capsid (left), with a zoomed-in P domain dimer outer-facing surface (middle) and a 90° dimer rotation that shows the difference in conservation of the outer- and inner-facing surfaces (right). The outer-facing surface (top portion) is substantially less conserved, with two major surface patches of conserved residues overlapping the HBGA binding site. (The highly conserved but nonprotruding portions of the capsid correspond to the shell domain.) (B) Close-up stereo view of panel A, middle, showing the six different HBGAs bound to the GII.10 P domain. (C) Surface representation of GII.10 amino acid conservation was obtained as described above and mapped onto the GII.10 P domain structure. The identified site of vulnerability (yellow) was defined as the surface area of the following GII.10 residues participating in conserved hydrogen-bonding interactions with αfucose1-2: Asn355, Arg356, Asp385 from one subunit, and Gly451 from the other subunit.

placement of the correct disaccharide unit was clear from other structures. We note, however, that the density observed for a βfucose variant of the H type 2-trisaccharide looked very good, indicating that βfucose is accommodated by the norovirus binding pocket, in addition to the standard αfucose.

One reason that the recognition of the HBGAs could be reduced to a single saccharide unit may relate to the avidity between norovirus and HBGAs on the cell surface. It is likely that HBGA affinity correlates with the number of saccharide units fixed in the norovirus-HBGA interaction, and in some cases, only a single fucose was fixed. The expected low affinity between a single fucose and a norovirus virion is unlikely to provide sufficient affinity for receptor or cofactor function; interactions between a number of cell-associated fucoses and multiple binding sites on the polyvalent norovirus capsid, however, might suffice. Similar avidity considerations have been observed with influenza, where relatively weak interactions with sialic acid are sufficient to serve as receptors (33). The

observed primary binding to αfucose along with avidity considerations open up a number of possibilities for norovirus entry: in addition to HBGAs, for example, the α1-2fucosylation of mucin (see references 32 and 42) may potentially allow mucin to act as a receptor or cofactor. Indeed, since the rarely detected GII.10 P domain bound a panel of HBGAs and the αfucose1-2 binding interface was similar to that of the dominant outbreak GII.4 strain, other receptors or cofactors may be important determinants for genotype prevalence and/or viral entry.

Our structural analysis strengthens the previous observation that GII noroviruses recognition of HBGAs requires the preservation of a conserved binding site across a dimer interface, which involves interactions with both the P1 and P2 subdomains (41). It has been previously suggested that the P2 subdomain is an insertion into P1 and may be the determinant of strain specificity due to its high variability and surface exposure (30). In contrast, P1 is more conserved and more internal (30), suggesting that its role as a specificity determinant may be diminished. Our structures, as well as previously determined GII.4 complex structures (4), indicate that HBGA binding involves important contacts with residues on a P1-interface loop (Fig. 3, 4, 6, and 7). These results indicate that, in addition to being partially responsible for homodimerization (30), the P1 subdomain plays a prominent role in recognizing HBGAs and thus may play a more prominent role in strain specificity than previously suggested.

Overall, the results provide a framework for understanding how requirements for HBGA interactions influence norovirus sequence conservation and lead to a highly conserved site on the outer surface of the capsid. This highly conserved site is a potential site of vulnerability for inhibition of virus entry. Whether small molecule competition with or antibody targeting to this conserved site allows for effective norovirus inhibition of entry remains to be seen. As we observe here, diversity in HBGA recognition (between different genotypes, different HBGAs, and different units of each HBGA) and reductions in required HBGA affinity (through avidity) provide a mechanism for viral reduction in the size of the conserved surface area while maintaining functional requirements for interactions with the host during entry.

ACKNOWLEDGMENTS

We thank the members of the Structural Biology Section and Structural Bioinformatics Core at the NIH Vaccine Research Center for comments on the manuscript, P. Adams and G. Murshudov for assistance with glycan refinement, K. Bok and K. Green for discussions on HBGA recognition, S. Hussan and C. A. Bewley for NMR observation of H type 2-trisaccharide, S. Park for assistance with Hiro refinement, and J. Stuckey for assistance with figures.

G.S.H., K.K., and P.D.K. designed the research; G.S.H., C.B., I.G., J.S.M., L.C., and T.Z. performed the research; G.S.H., C.B., I.G., J.S.M., K.K., and P.D.K. analyzed the data; and G.S.H., I.G., and P.D.K. wrote the first draft of the paper, on which all authors commented.

Support for this work was provided by the Intramural Research Program of the National Institutes of Health (United States), by a grant from The Japan Health Science Foundation, and by grants from the Ministry of Health, Labor, and Welfare of Japan. Use of sector 22 (Southeast Region Collaborative Access Team) at the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract W-31-109-Eng-38.

REFERENCES

1. **Adams, P. D., et al.** 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. **66:**213–221.
2. **Ahmad, S., M. Gromiha, H. Fawareh, and A. Sarai.** 2004. ASAView: database and tool for solvent accessibility representation in proteins. BMC Bioinformatics **5:**51.
3. **Bu, W., et al.** 2008. Structural basis for the receptor binding specificity of Norwalk virus. J. Virol. **82:**5340–5347.
4. **Cao, S., et al.** 2007. Structural basis for the recognition of blood group trisaccharides by norovirus. J. Virol. **81:**5949–5957.
5. **Carrillo-Tripp, M., et al.** 2009. VIPERdb2: an enhanced and web API enabled relational database for structural virology. Nucleic Acids Res. **37:** D436–D442.
6. **Choi, J. M., A. M. Hutson, M. K. Estes, and B. V. Prasad.** 2008. Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. Proc. Natl. Acad. Sci. U. S. A. **105:**9175–9180.
7. **Collaborative Computational Project, Number 4.** 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50:**760–763.
8. **Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan.** 2010. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. **66:**486–501.
9. **Hansman, G. S., et al.** 2004. Detection of norovirus and sapovirus infection among children with gastroenteritis in Ho Chi Minh City, Vietnam. Arch. Virol. **149:**1673–1688.
10. **Hansman, G. S., et al.** 2006. Genetic and antigenic diversity among noroviruses. J. Gen. Virol. **87:**909–919.
11. **Harrington, P. R., L. Lindesmith, B. Yount, C. L. Moe, and R. S. Baric.** 2002. Binding of Norwalk virus-like particles to ABH histo-blood group antigens is blocked by antisera from infected human volunteers or experimentally vaccinated mice. J. Virol. **76:**12335–12343.
12. **Harrington, P. R., J. Vinje, C. L. Moe, and R. S. Baric.** 2004. Norovirus capture with histo-blood group antigens reveals novel virus-ligand interactions. J. Virol. **78:**3035–3045.
13. **Huang, P., et al.** 2003. Noroviruses bind to human ABO, Lewis, and secretor histo-blood group antigens: identification of 4 distinct strain-specific patterns. J. Infect. Dis. **188:**19–31.
14. **Hutson, A. M., R. L. Atmar, and M. K. Estes.** 2004. Norovirus disease: changing epidemiology and host susceptibility factors. Trends Microbiol. **12:**279–287.
15. **Hutson, A. M., R. L. Atmar, D. Y. Graham, and M. K. Estes.** 2002. Norwalk virus infection and disease is associated with ABO histo-blood group type. J. Infect. Dis. **185:**1335–1337.
16. **Jiang, X., M. Wang, D. Y. Graham, and M. K. Estes.** 1992. Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. J. Virol. **66:**6527–6532.
17. **Kabsch, W.** 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Crystallogr. **26:**795–800.
18. **Kageyama, T., et al.** 2004. Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to Norovirus in Japan. J. Clin. Microbiol. **42:**2988–2995.
19. **Lindesmith, L., et al.** 2003. Human susceptibility and resistance to Norwalk virus infection. Nat. Med. **9:**548–553.
20. **Lindesmith, L. C., et al.** 2010. Heterotypic humoral and cellular immune responses following Norwalk virus infection. J. Virol. **84:**1800–1815.
21. **Lindesmith, L. C., et al.** 2008. Mechanisms of GII.4 norovirus persistence in human populations. PLoS Med. **5:**e31.
22. **Marionneau, S., et al.** 2001. ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. Biochimie **83:**565–573.
23. **Matsui, S. M., and H. B. Greenberg.** 2000. Immunity to calicivirus infection. J. Infect. Dis. **181**(Suppl. 2):S331–S315.
24. **McCoy, A. J., et al.** 2007. Phaser crystallographic software. J. Appl. Crystallogr. **40:**658–674.
25. **Noel, J. S., R. L. Fankhauser, T. Ando, S. S. Monroe, and R. I. Glass.** 1999. Identification of a distinct common strain of "Norwalk-like viruses" having a global distribution. J. Infect. Dis. **179:**1334–1344.
26. **Otwinowski, Z., and W. Minor.** 1997. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. **276:**307–326.
27. **Parrino, T. A., D. S. Schreiber, J. S. Trier, A. Z. Kapikian, and N. R. Blacklow.** 1977. Clinical immunity in acute gastroenteritis caused by Norwalk agent. N. Engl. J. Med. **297:**86–89.
28. **Pei, J., and N. V. Grishin.** 2001. AL2CO: calculation of positional conservation in a protein sequence alignment. Bioinformatics **17:**700–712.
29. **Pettersen, E. F., et al.** 2004. UCSF chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. **25:**1605–1612.
30. **Prasad, B. V., et al.** 1999. X-ray crystallographic structure of the Norwalk virus capsid. Science **286:**287–290.
31. **Rockx, B. H., H. Vennema, C. J. Hoebe, E. Duizer, and M. P. Koopmans.**

2005. Association of histo-blood group antigens and susceptibility to norovirus infections. J. Infect. Dis. **191:**749–754.
32. **Ruvoen-Clouet, N., et al.** 2006. Bile-salt-stimulated lipase and mucins from milk of 'secretor' mothers inhibit the binding of Norwalk virus capsids to their carbohydrate ligands. Biochem. J. **393:**627–634.
33. **Sauter, N. K., et al.** 1989. Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study. Biochemistry **28:**8388–8396.
34. **Shirato-Horikoshi, H., S. Ogawa, T. Wakita, N. Takeda, and G. S. Hansman.** 2007. Binding activity of norovirus and sapovirus to histo-blood group antigens. Arch. Virol. **152:**457–461.
35. **Siebenga, J. J., et al.** 2010. Phylodynamic reconstruction reveals norovirus GII.4 epidemic expansions and their molecular determinants. PLoS Pathog. **6:**e1000884.
36. **Siebenga, J. J., et al.** 2007. Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006. J. Virol. **81:**9932–9941.
37. **Tan, M., R. S. Hegde, and X. Jiang.** 2004. The p domain of norovirus capsid protein forms dimer and binds to histo-blood group antigen receptors. J. Virol. **78:**6233–6242.
38. **Tan, M., et al.** 2003. Mutations within the P2 domain of norovirus capsid affect binding to human histo-blood group antigens: evidence for a binding pocket. J. Virol. **77:**12562–12571.
39. **Tan, M., and J. X. Jiang.** 2010. Virus-host interaction and cellular receptors of caliciviruses, p. 111–129. *In* G. S. Hansman, J. X. Jiang, and K. Y. Green (ed.), Caliciviruses molecular and cellular virology, 1st ed. Caister Academic Press, Norwich, United Kingdom.
40. **Tan, M., and X. Jiang.** 2005. Norovirus and its histo-blood group antigen receptors: an answer to a historical puzzle. Trends Microbiol. **13:**285–293.
41. **Tan, M., et al.** 2009. Conservation of carbohydrate binding interfaces: evidence of human HBGA selection in norovirus evolution. PLoS One **4:**e5058.
42. **Tian, P., M. Brandl, and R. Mandrell.** 2005. Porcine gastric mucin binds to recombinant norovirus particles and competitively inhibits their binding to histo-blood group antigens and Caco-2 cells. Lett. Appl. Microbiol. **41:**315–320.
43. **Weis, W., et al.** 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. Nature **333:**426–431.
44. **Wiley, D. C., I. A. Wilson, and J. J. Skehel.** 1981. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature **289:**373–378.
45. **Wu, X., et al.** 2010. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science **329:**856–861.
46. **Wyatt, R. G., et al.** 1974. Comparison of three agents of acute infectious nonbacterial gastroenteritis by cross-challenge in volunteers. J. Infect. Dis. **129:**709–714.
47. **Zheng, D. P., et al.** 2006. Norovirus classification and proposed strain nomenclature. Virology **346:**312–323.
48. **Zhou, T., et al.** 2010. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. Science **329:**811–817.

JVI Journals.ASM.org

Structural Basis for Norovirus Inhibition and Fucose Mimicry by Citrate

Grant S. Hansman,[a,b] Syed Shahzad-ul-Hussan,[a,c] Jason S. McLellan,[a] Gwo-Yu Chuang,[a] Ivelin Georgiev,[a] Takashi Shimoike,[b] Kazuhiko Katayama,[b] Carole A. Bewley,[c] and Peter D. Kwong[a]

Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA[a]; Department of Virology II, National Institute of Infectious Diseases, Tokyo, 208-0011, Japan[b]; and Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA[c]

Human noroviruses bind with their capsid-protruding domains to histo-blood-group antigens (HBGAs), an interaction thought to direct their entry into cells. Although human noroviruses are the major cause of gastroenteritis outbreaks, development of antivirals has been lacking, mainly because human noroviruses cannot be cultivated. Here we use X-ray crystallography and saturation transfer difference nuclear magnetic resonance (STD NMR) to analyze the interaction of citrate with genogroup II (GII) noroviruses. Crystals of citrate in complex with the protruding domain from norovirus GII.10 Vietnam026 diffracted to 1.4 Å and showed a single citrate bound at the site of HBGA interaction. The citrate interaction was coordinated with a set of capsid interactions almost identical to that involved in recognizing the terminal HBGA fucose, the saccharide which forms the primary conserved interaction between HBGAs and GII noroviruses. Citrate and a water molecule formed a ring-like structure that mimicked the pyranoside ring of fucose. STD NMR showed the protruding domain to have weak affinity for citrate (460 μM). This affinity, however, was similar to the affinities of the protruding domain for fucose (460 μM) and H type 2 trisaccharide (390 μM), an HBGA shown previously to be specifically recognized by human noroviruses. Importantly, competition STD NMR showed that citrate could compete with HBGA for norovirus binding. Together, the results suggest that citrate and other glycomimetics have the potential to block human noroviruses from binding to HBGAs.

Human noroviruses, family *Caliciviridae*, are the dominant cause of outbreaks of gastroenteritis. Many aspects of human norovirus replication, however, remain unclear, mainly because these viruses cannot be grown in cell culture. Transmission predominately occurs through ingestion of contaminated foods, airborne transmission, and person-to-person contact. Medical treatment usually involves orally administered fluids and electrolyte replacement therapy. Currently, there is no effective vaccine.

Human noroviruses can be divided into 2 main genogroups (GI and GII), which can be further subdivided into at least 25 different genotypes (GI.1 to -8 and GII.1 to -17) (26, 57). The norovirus genome has three open reading frames (ORFs) that encode nonstructural, capsid, and small structural proteins, respectively. The capsid of human norovirus is composed of two domains, shell and protruding (P) domains. The shell forms a scaffold around the RNA, and the dimeric P domain contains determinants for both antigenicity and receptor binding (25, 43, 51). The P domain is further subdivided into P1 and P2 subdomains, where the P1 subdomain interacts with the shell domain and is buried under the outermost P2 subdomain.

Human noroviruses bind to histo-blood group antigens (HBGAs), with recognition occurring in the P domain. HBGAs are complex carbohydrates present on mucosal epithelial cells or free antigens in blood, saliva, and other fluids (32). X-ray crystal structures of norovirus P domains in complex with different HBGAs have defined distinct binding sites for GI and GII viruses (8, 11, 12, 21); in particular, the HBGA binding site of GII is located at the dimeric interface of two P domains, whereas the HBGA binding site in GI is located within a single P domain (8, 11, 12, 21).

A number of recent studies have shown that natural fruits or their constituents, including orange juice, pomegranate juice, cranberry juice, and grape seed extract, can inhibit and/or reduce feline calicivirus and murine norovirus infectivity (23, 48–50, 54). Although there have been no studies to support the idea that natural fruits or their constituents can prevent human norovirus infections, and data on the mode of inhibition of fruits have been lacking, the stability of human norovirus virus-like particles over a pH range of 3 to 7 (3) suggested that the effect might be related to a specific interaction with compounds in fruits rather than a pH effect. In this study, we used X-ray crystallography and saturation transfer difference nuclear magnetic resonance (STD NMR) to provide atomic-level details on the interaction of citrate and GII human noroviruses. We show that citrate specifically binds at the HBGA recognition site of GII noroviruses, and this inhibits P domain binding of both fucose and HBGA.

MATERIALS AND METHODS

Protein expression, purification, and crystallization of the norovirus P domain. The norovirus Vietnam026 GII.10 P domain (GenBank accession no. AF504671) (22) was expressed in *Escherichia coli* as previously

Received 5 August 2011 Accepted 17 October 2011
Published ahead of print 26 October 2011
Address correspondence to Peter D. Kwong, pkwong@mail.nih.gov, or Carole A. Bewley, caroleb@mail.nih.gov.
G.S.H. and S.S.H. contributed equally to this article.
Supplemental material for this article may be found at http://jvi.asm.org/.

doi:10.1128/JVI.05909-11

described (21). Briefly, a truncated form of the GII.10 P domain was optimized for *E. coli* expression, cloned in a modified pMal-c2x vector at the BamHI and NotI sites (New England BioLabs), and transformed into *E. coli* BL21 cells (Invitrogen), and expression was induced with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) for 18 h at 22°C. After a series of purifications and cleavage steps, the P domain was concentrated to 2 mg/ml and stored in gel filtration buffer (0.35 M NaCl, 2.5 mM Tris [pH 7.0], 0.02% NaN_3) before crystallization. Crystals of the P domain were obtained by the hanging-drop vapor diffusion method, with the mother solution containing citric acid triammonium (0.66 M [pH 6.5]) and isopropanol (1.65% [vol/vol]).

Data collection, structure solution, and refinement. X-ray diffraction data at a 1.000-Å wavelength were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-BM at the Advanced Photon Source, Argonne National Laboratory, Argonne, IL, and processed with HKL2000 (41). Prior to data collection, crystals were transferred to a cryoprotectant solution consisting of the mother liquor in 30% ethylene glycol, loop mounted, and flash-cooled in a nitrogen cryostat to 100°K. Structures were solved by molecular replacement with PHASER (35) by using Protein Data Bank (PDB) code 2OBR (11) as a search model. Structures were refined in multiple rounds of manual model building in COOT (16) and positional together with TLS refinement in REFMAC (13) and PHENIX (1).

Structure analysis and figures. Citrate and H type 2 interactions were determined using Discovery Studio (Accelrys, version v2.5.5.9350). Figures were rendered using PyMOL (Schroedinger, LLC, version 1.2r3) and ChemDraw Ultra (Cambridgesoft, version 12.0.2.1076).

STD NMR. All NMR data were recorded at 298°K on a Bruker Avance 600 NMR spectrometer equipped with a cryogenically cooled *z*-shielded gradient probe. One-dimensional (1D) STD NMR spectra were acquired with selective irradiation at −1 and +40 ppm (on and off resonance, respectively) using a train of 50-ms Gaussian-shaped radio frequency pulses separated by 1-ms delays and an optimized power level of 57 db. Water suppression was achieved with a binomial 3-9-19 pulse sequence. Samples were prepared in 20 mM sodium phosphate buffer containing 50 mM sodium chloride at pH 6.8. The NMR data were processed and analyzed with Topspin 2.1. STD enhancements were expressed as the STD amplification factor, A_{STD}, defined as $A_{STD} = (I_0 - I_{SAT})\, I_0^{-1}\, ([L_t]/[P])$, where L_t and P are the total ligand and protein concentrations, respectively (34). HBGAs, H type 2 disaccharide [α-L-fucose-(1-2)-β-D-galactose], and H type 2 trisaccharide [α-L-fucose-(1-2)-β-D-galactose-(1-4)-2-*N*-acetyl-β-D-glucosamine] were purchased from V-labs, and L-fucose was purchased from Sigma-Aldrich. For the citrate experiments, sodium citrate dihydrate (Sigma-Aldrich) was added to sodium phosphate buffer and then titrated at pH 6.85 ± 0.1.

Computational citrate docking studies of other saccharide-binding proteins. Citrate docking analyses were performed against six different saccharide-binding proteins, including *Anguilla anguilla* agglutinin (PDB identification no. 1K12) (5), *Aleuria aurantia* lectin (PDB identification no. 1IUC) (19), *Streptococcus pneumoniae* virulence factor SpGH98 (PDB identification no. 2J1S) (7), *Pseudomonas aeruginosa* PA-IIL lectin (PDB identification no. 2JDH) (33), parainfluenza virus 5 hemagglutinin-neuraminidase (PDB identification no. 1Z4X) (56), and porcine adenovirus type 4 galectin domain (PDB identification no. 2WSV) (20). Water molecules and ligands were removed from the PDB files, with the exception of one water molecule (HOH 935) in 1Z4X, which is present in both ligand-free and sialyllactose-bound hemagglutinin-neuraminidase structures. For 2JDH, the two calcium ions in the fucose binding site were kept, and the partial charges for the calcium ions were assigned to 1.5 as suggested by previous studies (38). AutoDock4.2 (39) was used as the docking engine, with the grid files generated by Autogrid4.2 using default parameters and centered on the cocrystallized ligands. The citrate molecule was docked to the three structures using default parameters (ga_pop_size = 150, ga_num_evals = 2,500,000, and ga_run = 50). For each structure, the docking pose with the lowest estimated free energy of binding among the 50 docking runs was selected as the predicted binding pose. For comparison, for each complex, the cocrystallized ligand (or the terminal monosaccharide having the largest contact area with the binding site, if the cocrystallized ligand was not a monosaccharide) was docked in the saccharide binding site using the same procedure. Fucose and citrate molecules were also docked to the fucose-bound GII.10 P domain (PDB identification no. 3ONY) and the citrate-bound GII.10 P domain, respectively, for comparison. The water molecule (HOH 135) mediating the interaction between citrate and the protein was present during the citrate docking analysis.

Protein structure accession number. Atomic coordinate and structure factors for the citrate-bound GII.10 P domain have been deposited in the Protein Data Bank under accession no. 3RY8.

RESULTS

X-ray crystal structure of citrate bound to the GII.10 P domain. The GII.10 P domain protein could be expressed in *E. coli* to 2 mg/liter and was purified and prepared for crystallization as previously described (21). To obtain a GII.10 P domain-citrate complex, we chose a crystallization condition that was similar to our previously reported GII.10 P domain-HBGA complex conditions (21), though with the addition of citrate. The GII.10 P domain-citrate complex formed rectangular plate crystals, and X-ray diffraction data revealed a space group of $P2_1$, the same as the previous GII.10 P domain-HBGA complexes (21), and strong diffraction to 1.4 Å. Structure solution by molecular replacement revealed one dimer per asymmetrical unit (Fig. 1A), and refinement led to an R_{value} of 0.139 (R_{free} = 0.151), with well-defined density for most of the P domain dimer (Table 1). Electron density for residues 296 to 299 (chain A) and 296 to 300 and 344 to 351 (chain B) was poor, and these residues were not modeled. Extra electron density was observed at the HBGA binding site, where a single citrate molecule was clearly distinguished and refined (Fig. 1B; see Fig. S1 in the supplemental material). The structure of the GII.10 P domain in complex with citrate was highly reminiscent of the other known structures (GI.1, GII.4, and GII.12), where the P1 subdomain contains a single α-helix and the P2 subdomain contains six anti-parallel β-strands that form two anti-parallel β-sheets (21).

Citrate was highly coordinated by the GII.10 P domain. At a 1.4-Å resolution, detailed interactions between citrate and the P domain could be defined. Seven residues of the P domain, many of which are conserved and located at the dimer interface, are involved in hydrogen bonding interactions with citrate (Fig. 1B and C). These include the side chain of Tyr452 and main chain of Gly451 from one P domain subunit as well as side chains Arg356 and Asp385 and the main chain of Asn355 of a second P domain subunit. Unique to citrate binding, side chains of Asn342 and Ser387 make a water-mediated hydrogen bond with the C-5 CO group of citrate. Superposition of citrate-bound and apo GII.10 P domain structures indicated that the citrate interaction did not cause any conformation changes in the GII.10 P domain.

Comparisons of citrate and HBGA interaction with the GII.10 P domain. Compared with GII.10 P domains in complex with HBGAs, we found that citrate essentially mimics the fucose unit of HBGAs. By using H type 2 di- and trisaccharides as examples, superposition of the P domains revealed that three carbon atoms, including C-2, C-3, and the C-3 carboxy carbon, and three oxygen atoms, including the C-1 and C-3 carboxy oxygens and the C-3 hydroxyl group of citrate closely overlapped with C-5/C-4/C-3/O-5/O-4/O-3 of the terminal α-fucose ring (Fig. 2). In addition,

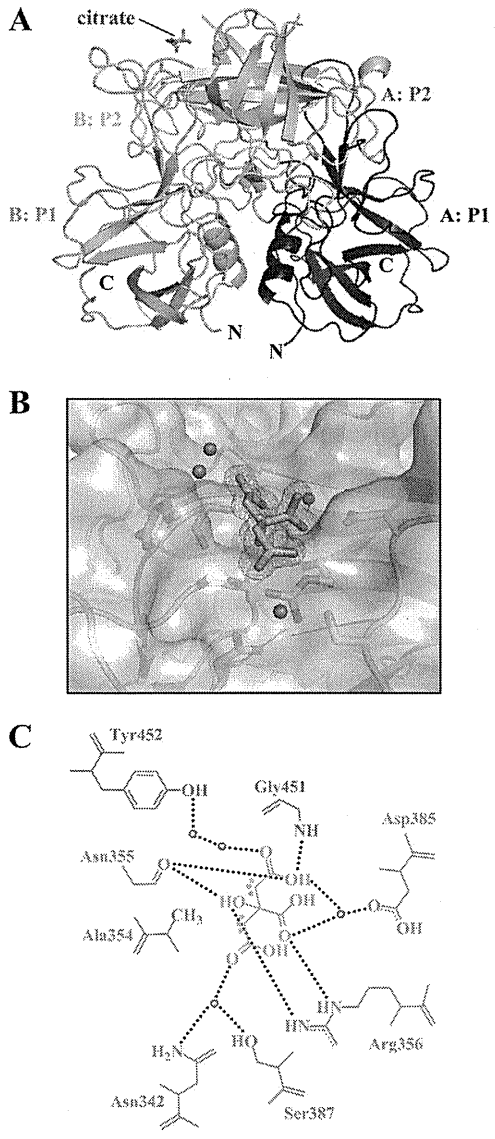

FIG 1 Citrate binding to the GII.10 P domain. (A) X-ray crystal structure of the GII.10 P domain dimer (ribbon structure) and the bound citrate (green sticks). Each P subdomain is colored differently: i.e., chain A, P1, blue; chain A, P2, light blue; chain B, P1, violet; and chain B, P2, salmon. (B) Surface representation of the GII.10 P domain (colored as in panel A) showing the residues (sticks) and water molecules (red spheres) interacting with the citrate molecule (green sticks). The $2F_o$-F_c density was contoured at 1.0 σ. (C) Residues interacting with the citrate molecule were contributed by both monomers (colored as in panel A), where the black dotted lines represent the hydrogen bonds, the cyan dots near the citrate represent the hydrophobic interactions with Ala354, and the red spheres represent water molecules. For simplicity, only the backbone is shown for residues that were backbone mediated. Hydrogen bond distances were less than 3.1 Å, though the majority were ~2.8 Å.

a water molecule, present in the citrate-bound structure but absent from the HBGA-bound structures, occupied the site of the C-2 hydroxyl of fucose (Fig. 2). In this configuration, the citrate and associated water molecule formed a ring-like structure, mimicking the pyranoside ring of fucose. Finally, the comparisons showed that of the seven residues involved in hydrogen bonding interactions with citrate, five made almost identical interactions with their comparable atoms in fucose.

Characterization of binding of citrate, H type 2 trisaccharide, and fucose to GII.10 P domain by STD NMR. Given the remarkable similarities observed for citrate and fucose binding to the GII.10 P domain by crystallography, we sought to characterize in solution by NMR the binding of GII.10 P domain with citrate, HBGAs, and fucose and ultimately to determine their relative binding affinities and whether they bind competitively.

STD enhancements were observed for methylene protons H2A and H2B of citrate, consistent with their close proximity to the protein in the bound state (Fig. 3A). In the crystal structure, these hydrogens are within van der Waals contact of the methyl of Ala354 (Fig. 1C). With H type 2 trisaccharide, the most prominent STD signals that could be assigned corresponded to H-1, H-2, and H-4 of α-fucose; H-3 of galactose; and H-1, H-2, and *N*-acetyl of glucosamine (Fig. 3B). We also characterized binding of monosaccharide α/β-fucopyranose, as it also would be used in competition STD NMR experiments. As seen in Fig. 3C, binding of both anomers was observed, with H-1, H-2, and H-4 of α-fucose versus H-2, H-4, and H-5 of β-fucose showing the strongest enhancements. Although natural H type 2 HBGAs contain α-Fuc(1-2)Gal and not β-Fuc, it is interesting that the HBGA binding site of norovirus can bind both. By NMR, we observed binding of α and β forms of the monosaccharide (Fig. 3C) as well as synthetic H type 2 trisaccharide α/β-Fuc(1-2)β-Gal(1-4)β-GlcN (Fig. 3B) and H type 2 disaccharide α-Fuc(1-2)β-Gal(1-4) (data not shown), and by crystallography, binding of synthetic H type 2

TABLE 1 Data collection and refinement statistics for structures of the GII.10 Vietnam026 norovirus P domain[a]

Parameter	Value(s) for citrate (026_citrate; PDB accession no. 3RY8)[b]
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	63.76, 79.81, 69.60
α, β, γ (°)	90, 96.84, 90
Resolution (Å)	50-1.40 (1.45-1.40)
R_{sym}	7.3 (30.5)
$I/\sigma I$	18.7 (3.2)
Completeness (%)	99.9 (99.6)
Redundancy	3.7 (3.2)
Refinement	
Resolution range (Å)	31.98-1.399
No. of reflections	131,576
R_{work}/R_{free}	0.1388/0.1506
No. of atoms:	
Total	5,587
Protein	4,722
Ligand/ion	57
Water	808
B-factors	
Protein	18.3
Ligand/ion	19.2
Water	30.2
RMSD	
Bond length (Å)	0.011
Bond angle (°)	1.393

[a] Each data set was collected from a single crystal.
[b] Values in parentheses are for the highest-resolution shell.

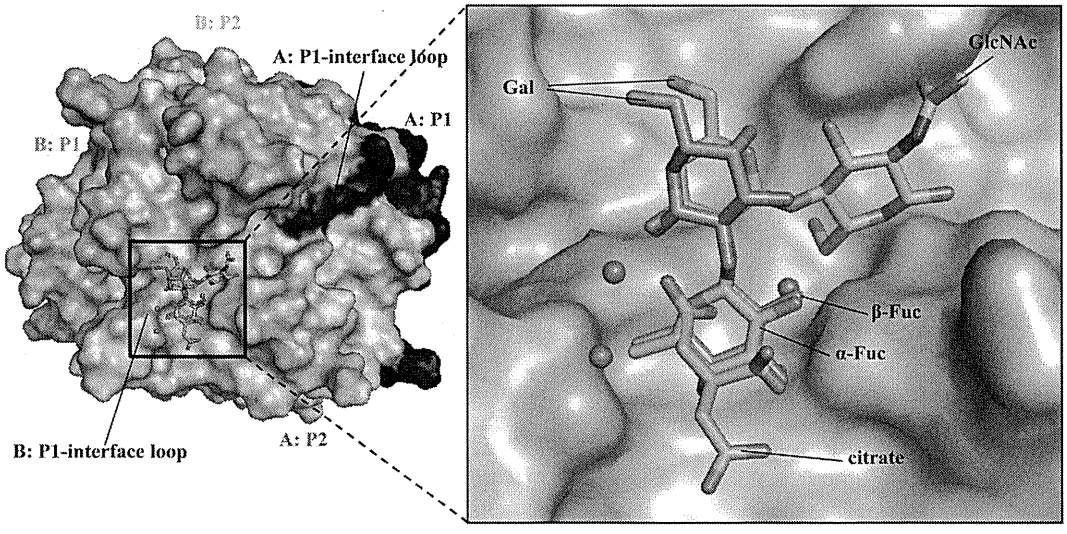

FIG 2 The HBGA and citrate binding site overlapped on the top of the GII.10 P domain. (A) The citrate molecule (green sticks) bound at the HBGA binding site; shown here are the bound H type 2 di- and trisaccharides (orange and cyan sticks, respectively). (B) Close-up of the black square in panel A, showing the H type 2 disaccharide [α-L-fucose(1-2)-β-D-galactose] and H type 2 trisaccharide [β-L-fucose(1-2)-β-D-galactose(1-4)-2-*N*-acetyl-β-D-glucosamine].

trisaccharide β-Fuc(1-2)β-Gal(1-4)β-GlcN was observed, in addition to binding of other HBGAs containing the α-Fuc(1-2)β-Gal linkage (21). Finally, it is interesting to note that a similar mode of citrate binding was observed for the soluble GII.12 P domain (see Fig. S2 in the supplemental material).

Affinity of citrate, H type 2 tri- and disaccharides, and L-fucose to the GII.10 P domain. We used single-ligand titration STD NMR experiments to determine the K_D (equilibrium dissociation constant) of citrate binding to the GII.10 P domain (Fig. 4A) (2). STD amplification factors (A_{STD}) (34) were calculated by integrating the signal at δ_H 2.54 ppm in difference and corresponding reference spectra. Initial growth rates (A_{0STD}) were obtained by measuring the effect on A_{STD} as a function of various saturation time (t_{sat}) and fitting the data to the equation $A_{STD} = A_{max\ STD}[1 - \exp(-kt_{sat})]$ for each concentration (300, 600, 900, 1,200, and 1,500 μM) of the ligand. The K_D of citrate was in turn measured as 460 ± 80 μM by fitting A_{0STD} values as a function of ligand concentration using the equation $y = B_{max}/(K_D + x)$, where x is the ligand concentration and B_{max} represents the plateau of the curve (Fig. 4) (2, 37). For H type 2 trisaccharide, the STD enhancements for the *N*-acetyl signal were sufficiently strong to allow for accurate integration, even at very short saturation times (0.1 s); thus, a K_D value of 390 ± 70 μM could be determined directly by fitting A_{STD} values as a function of ligand concentration (40) (Fig. 5). K_D values for fucose and H type 2 disaccharides were in turn obtained from single point competition STD experiments as described previously (36) to give values of 460 ± 10 and 420 ± 40 μM, respectively (see Fig. S3 in the supplemental material).

Competition of HBGAs and citrate with the GII.10 P domain. To confirm the overlapping mode of binding observed in the crystal structure of citrate and fucose of H type 2 ligands, A_{STD} values of L-fucopyranose and H type 2 trisaccharide were monitored while titrating citrate to the samples. As seen in Fig. 4E, addition of citrate to a sample of P domain-H type 2 trisaccharide diminishes the trisaccharide signals in a concentration-dependent manner, indicating that citrate directly competes with the trisaccharide for P domain binding, giving a K_i of 600 ± 20 μM. Upon addition of citrate, the pH of the solutions was found to remain constant (pH 7.2 ± 0.1), indicating that the competition was a direct result of citrate binding rather than pH. The same effect was observed in STD competition experiments with L-fucose (data not shown). Importantly, the reverse set of experiments showed that HBGAs can compete with citrate for P domain binding (data not shown), indicating that the P domain is unaffected by the presence of citrate. Together, these results conclusively demonstrate molecular mimicry between citrate and fucose of HBGAs.

DISCUSSION

Despite the discovery of human norovirus nearly 40 years ago (27), little is known about the capsid interaction with ligands (18, 44) other than HBGAs (8, 11, 12, 15, 21, 45). Our finding that citrate binds at the terminal fucose binding site was somewhat unexpected, given that the structure of citrate is unlike the structure of fucose and considering that the GII.10 P domain could not bind HBGAs having an α-fucose1-3/4 saccharide (21). In an earlier enzyme immune assay study, Feng et al. screened ~5,000 compounds (the Diversity screening set; Timtec, Inc.) for their ability to block GI and GII norovirus virus-like particles (VLPs) from binding to saliva samples of known HBGA type (18). They found 14 compounds that had strong inhibition; however, the mode of action was not determined. In a more recent NMR study, Rademacher et al. screened ~500 compounds (the Maybridge Ro5 fragment library; Thermo Fisher Scientific, Inc.) for their ability to bind to a GII.4 VLP HBGA binding site (44). They showed that both univalent and multivalent compounds were capable of binding to the HBGA binding site. Interestingly, for both studies, the compounds that showed the highest affinities included compounds with at least one ring component. Taken together, these studies indicated that the HBGA binding site was capable of binding numerous compounds other than HBGAs, ranging from the small (smallest) citrate molecule to larger multivalent compounds.

For over a decade, the GII.4 noroviruses have remained as the dominant genotype of outbreaks of gastroenteritis around the

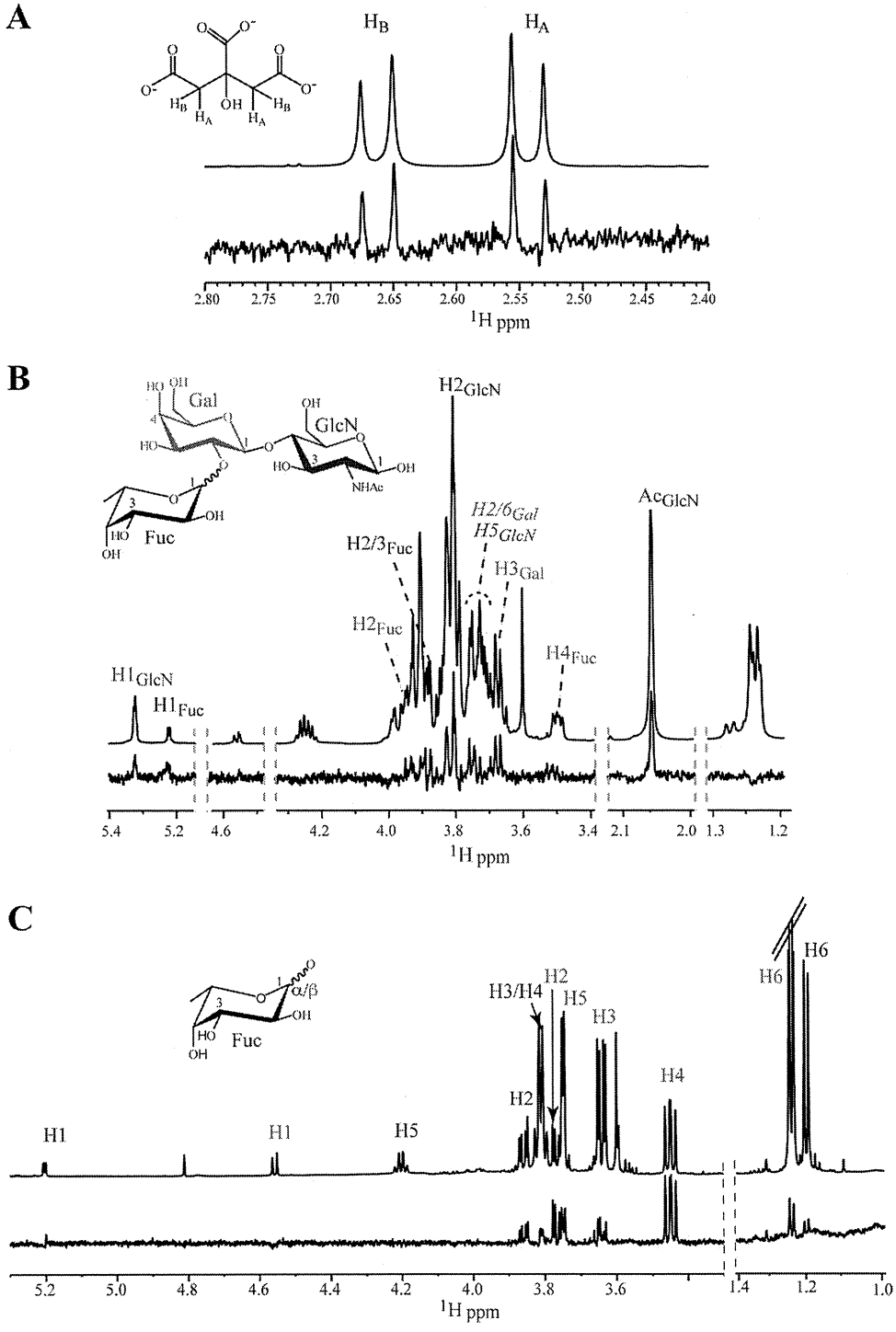

FIG 3 STD NMR spectra for citrate, H type 2 trisaccharide, and L-fucose bound to the GII.10 P domain. STD (lower) and reference (upper) spectra of (A) citrate (1.2 mM), (B) H type 2 trisaccharide (1.2 mM), and (C) L-fucose (mixture of α and β anomers) (1.2 mM) in the presence of the GII.10 P domain (15 μM). Nonoverlapping protons that exhibit STD enhancements are labeled and color coded by sugar residue, and signals for β-Fuc are red. One group of overlapping signals appears in italics.

world and as such the most well studied. Most studies agreed that a dominant GII.4 norovirus was replaced the following year or next by a new GII.4 "variant" norovirus that had ~5% amino acid change in the capsid gene (6, 9, 10, 30, 31, 47). The reason that the GII.4 variants dominated and not some other genotype was unknown, but studies have shown specific mutations at or surrounding the HBGA binding site were capable of altering the HBGA binding patterns (15, 30, 31, 52). These small changes were thought to lead to new GII.4 variants capable of causing pandemics, analogous to influenza A virus evolution (14, 29). Despite

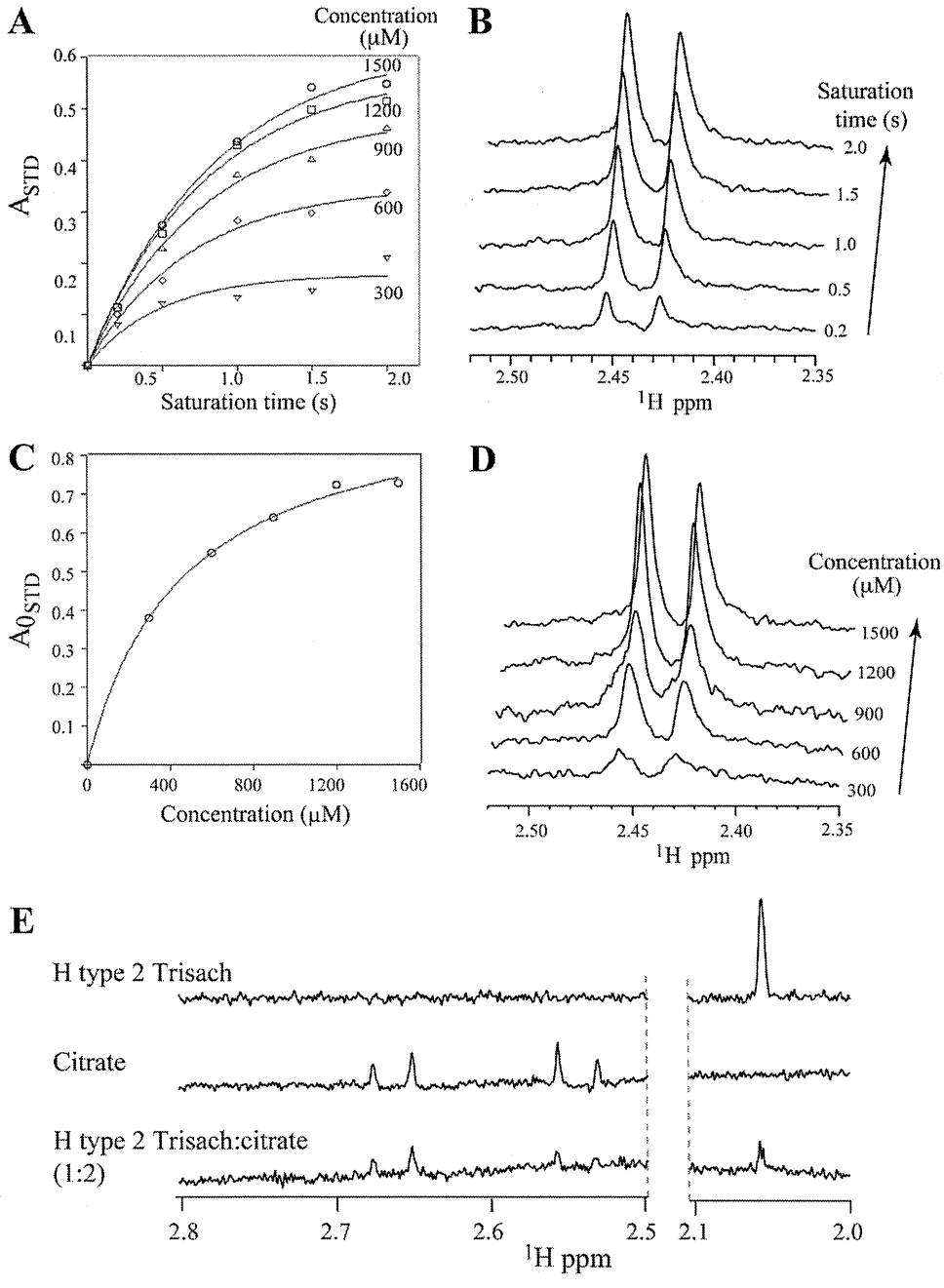

FIG 4 Binding affinity of citrate and HBGAs to GII.10 P domain by STD NMR. Data were used to obtain the K_D for citrate and H type 2 trisaccharide (Trisach) binding to GII.10 P by single-ligand titration STD NMR experiments (2). (A) Effect on STD enhancement (expressed as A_{STD}) (34, 37) as a function of saturation time (t_{sat}) and ligand concentration; (B) stacked plots of spectra for 1.5 mM citrate as a function of t_{sat} (y axis); (C) Langmuir binding curve used to obtain the K_D from the initial slope of A_{STD} as a function of citrate concentration. (D) Stacked plot of various citrate concentrations (t_{sat}, 2 s, 15 μM protein); (E) competition STD spectra of H type 2 trisaccharide (top), citrate (middle), and 1:2 H type 2 trisaccharide-citrate (0.75:1.5 mM; bottom) used to calculate the K_i of citrate (36).

these amino acid changes, few if any occurred at the fucose-binding site, thus highlighting the common site of vulnerability for GII noroviruses, especially for the pandemic GII.4 variant noroviruses. It is not known if the GI noroviruses will bind citrate given that the GI and GII P domain interactions with HBGAs were different, but since GI.1 P domain interacted with α-fucose1-2 and it was reported that the HBGA binding site was conserved among GI noroviruses (12), we suspect that GI noroviruses may also bind citrate, although further structural studies are needed.

Our unexpected finding that citrate and fucose have similar binding modes to the norovirus GII.10 P domain raises the question of whether such citrate mimicry of monosaccharide binding could be a general phenomenon or whether it is spe-

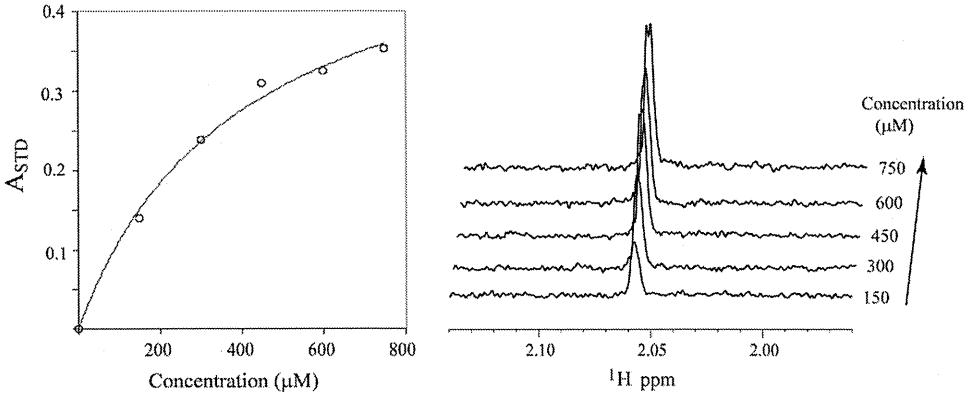

FIG 5 Binding affinity of H type 2 trisaccharide to the GII.10 P domain (left) effect on STD enhancement, expressed as A_{STD}, as a function of trisaccharide concentration in the presence of 15 μM GII.10 P domain. t_{sat} = 0.1 s. Curve fitting (described in the text) provides the K_D value. (Right) Stacked STD NMR spectra showing the change in enhancement of the nonoverlapped *N*-acetyl proton signals as a function of increasing concentration of H type 2 trisaccharide (40).

cific to norovirus and other caliciviruses. To investigate this, we performed *in silico* docking studies of citrate against four different fucose-binding proteins (*Anguilla anguilla* agglutinin, *Aleuria aurantia* lectin, *Streptococcus pneumoniae* virulence factor SpGH98, and *Pseudomonas aeruginosa* PA-IIL lectin) and two other saccharide-binding proteins (parainfluenza virus 5 hemagglutinin-neuraminidase and porcine adenovirus type 4 galectin domain), for which fucose or other saccharide-bound crystal structures were available (see Table S1 in the supplemental material). Computational docking analyses reveal different levels of citrate mimicry of monosaccharide binding for other saccharide-binding proteins. For *Anguilla anguilla* agglutinin, citrate, in its predicted binding pose, overlapped with the C-5, C-4, C-3, O-5, O-4, and O-3 atoms of fucose in a similar way to what was observed in the GII.10 P domain (Table S1), while forming hydrogen bonds with the same sets of protein residues as fucose (see Fig. S4 in the supplemental material). Citrate was thus predicted to show a high degree of mimicry to fucose, similarly to our experimental findings for the GII.10 P domain. For the other three fucose-binding proteins, citrate, in its predicted binding poses, did not overlap with the cocrystallized fucose, although it still formed the same sets of polar interactions as the cocrystallized fucose (see Fig. S5 to S7 in the supplemental material). Hence, our docking studies suggest that the mimicry between citrate and fucose binding observed for the GII.10 P domain could be a common, although not universal, phenomenon across other fucose-binding proteins. For all six fucose- and other saccharide-bound proteins for which docking was performed, the predicted citrate binding poses were able to form polar interactions with the same sets of protein residues as the cocrystallized ligand see (Fig. S4 to S9 in the supplemental material), indicating that citrate might be generally useful as a scaffold for designing glycomimetic inhibitors against these and other saccharide-interacting pathogens. Furthermore, a search of the ZINC database (4) revealed that there are more than three thousand compounds with at least 50% similarity to citrate. Thus, *in silico* screening of this database may present a promising approach for identifying small molecules that bind to saccharide-binding proteins. We note, however, that the predicted binding pose of citrate docked to fucose-bound GII.10 P domain had a root mean square deviation (RMSD) of 3.60 Å, while the predicted binding pose of citrate docked to citrate-bound GII.10 P domain with the cocrystallized water molecule had an RMSD of 1.87 Å. This indicates that the resulting docking modes could be error prone. Given that calculating small molecule-receptor binding energies is a difficult and error-prone task (24, 46), ultimately experimental validation would be necessary to confirm the generality of the citrate-saccharide mimicry predicted here.

The STD NMR data provided strong evidence that the integrity of the GII.10 P domain remained unchanged in the presence of different concentrations of citrate buffer and since the pH of the citrate buffer remained more or less the same during the titration, a specific effect of citrate was responsible for the reduction in HBGA attachment. Although the K_D values of citrate and H type 2 trisaccharide for the GII.10 P domain are in the range of 360 to 490 μM, these relatively weak affinities are typical for univalent protein-carbohydrate interactions (17, 28). Given that 90 copies of dimeric P domains are present on norovirus capsid, it is plausible that a multivalent version of citrate- or fucose-like ligands would greatly enhance affinities and provide a starting point for norovirus inhibitors. Indeed, Rademacher et al. show that multivalent fucose-like compounds have increased avidity over their univalent counterparts (44).

In conclusion, we have described the structural basis by which citrate binds to the HBGA binding site of the norovirus GII.10 P domain and can in turn inhibit HBGA binding. Natural compounds, such as juice from lemons and limes, which contain ~300 mM citric acid (42), may already reduce or inhibit norovirus infections, as suggested by a number of recent studies (23, 48–50, 54). In regard to this, it is tempting to speculate that a few drops of lemon juice with one's oysters might reduce norovirus infection. Epidemiological studies on the ingestion of foods high in citrate and norovirus infection may be illuminating, as may be correlations with related glycomimetics—e.g., with ascorbic acid (vitamin C). Controlled possibly volunteer studies should also provide an accurate assessment of norovirus inhibition. Additional compound screening will likely be required to identify a universal norovirus inhibitor with high potency and broad reactivity, and the structural basis for norovirus interaction with citrate as revealed here may be helpful in such efforts.

ACKNOWLEDGMENTS

We thank members of the Structural Biology Section and Structural Bioinformatics Core at the NIH Vaccine Research Center for comments on the manuscript and J. Stuckey for assistance with figures.

G.S.H., S.H., K.K., C.A.B., and P.D.K. designed the research, G.S.H., S.H., J.S.M., and T.S. performed the research, G.S.H., S.H., J.S.M., G.-Y.C., I.G., C.A.B., and P.D.K. analyzed the data, and G.S.H., S.H., G.-Y.C., I.G., C.A.B., and P.D.K. wrote the paper, on which all authors commented.

Support for this work was provided by the Intramural Research Program of the National Institutes of Health (NIDDK, C.A.B; NIAID, P.D.K); the Intramural AIDS Targeted Antiviral Program, Office of the Director, NIH (C.A.B.); a grant from The Japan Health Science Foundation; and by grants from the Ministry of Health, Labor, and Welfare of Japan. Use of sector 22 (Southeast Region Collaborative Access team) at the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract no. W-31-109-Eng-38.

REFERENCES

1. **Adams PD, et al.** 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. **66**:213–221.
2. **Angulo J, Enriquez-Navas PM, Nieto PM.** 2010. Ligand-receptor binding affinities from saturation transfer difference (STD) NMR spectroscopy: the binding isotherm of STD initial growth rates. Chemistry **16**:7803–7812.
3. **Ausar SF, Foubert TR, Hudson MH, Vedvick TS, Middaugh CR.** 2006. Conformational stability and disassembly of Norwalk virus-like particles. Effect of pH and temperature. J. Biol. Chem. **281**:19478–19488.
4. **Berman HM, et al.** 2000. The Protein Data Bank. Nucleic Acids Res. **28**:235–242.
5. **Bianchet MA, Odom EW, Vasta GR, Amzel LM.** 2002. A novel fucose recognition fold involved in innate immunity. Nat. Struct. Biol. **9**:628–634.
6. **Bok K, et al.** 2009. Evolutionary dynamics of GII. 4 noroviruses over a 34-year period. J. Virol. **83**:11890–11901.
7. **Boraston AB, Wang D, Burke RD.** 2006. Blood group antigen recognition by a Streptococcus pneumoniae virulence factor. J. Biol. Chem. **281**: 35263–35271.
8. **Bu W, et al.** 2008. Structural basis for the receptor binding specificity of Norwalk virus. J. Virol. **82**:5340–5347.
9. **Bull RA, Tu ET, McIver CJ, Rawlinson WD, White PA.** 2006. Emergence of a new norovirus genotype II.4 variant associated with global outbreaks of gastroenteritis. J. Clin. Microbiol. **44**:327–333.
10. **Bull RA, White PA.** 2011. Mechanisms of GII.4 norovirus evolution. Trends Microbiol. **19**:233–240.
11. **Cao S, et al.** 2007. Structural basis for the recognition of blood group trisaccharides by norovirus. J. Virol. **81**:5949–5957.
12. **Choi JM, Hutson AM, Estes MK, Prasad BV.** 2008. Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. Proc. Natl. Acad. Sci. U. S. A. **105**:9175–9180.
13. **Collaborative Computational Project N.** 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D. Biol. Crystallogr. **50**:760–763.
14. **De Jong JC, Rimmelzwaan GF, Fouchier RA, Osterhaus AD.** 2000. Influenza virus: a master of metamorphosis. J. Infect. **40**:218–228.
15. **de Rougemont A, et al.** 2011. Qualitative and quantitative analysis of the binding of GII.4 norovirus variants onto human blood group antigens. J. Virol. **85**:4057–4070.
16. **Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of Coot. Acta Crystallogr. D. Biol. Crystallogr. **66**:486–501.
17. **Fang M, Agha S, Lee R, Culpepper-Morgan J, D'Souza A.** 2000. Perforation of jejunal diverticulum: case report and review of literature. Conn. Med. **64**:7–10.
18. **Feng X, Jiang X.** 2006. Library screen for inhibitors targeting norovirus binding to their histo-blood group antigen receptors. Antimicrob. Agents Chemother. **51**:324–331.
19. **Fujihashi M, Peapus DH, Kamiya N, Nagata Y, Miki K.** 2003. Crystal structure of fucose-specific lectin from Aleuria aurantia binding ligands at three of its five sugar recognition sites. Biochemistry **42**:11093–11099.
20. **Guardado-Calvo P, et al.** 2010. Crystallographic structure of porcine adenovirus type 4 fiber head and galectin domains. J. Virol. **84**: 10558–10568.
21. **Hansman GS, et al.** 2011. Crystal structures of GII.10 and GII.12 norovirus protruding domains in complex with histo-blood group antigens reveal details for a potential site of vulnerability. J. Virol. **85**:6687–6701.
22. **Hansman GS, et al.** 2004. Detection of norovirus and sapovirus infection among children with gastroenteritis in Ho Chi Minh City, Vietnam. Arch. Virol. **149**:1673–1688.
23. **Horm KM, D'Souza DH.** 2011. Survival of human norovirus surrogates in milk, orange, and pomegranate juice, and juice blends at refrigeration (4 degrees C). Food Microbiol. **28**:1054–1061.
24. **Huey R, Morris GM, Olson AJ, Goodsell DS.** 2007. A semiempirical free energy force field with charge-based desolvation. J. Comput. Chem. **28**: 1145–1152.
25. **Jiang X, Wang M, Graham DY, Estes MK.** 1992. Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. J. Virol. **66**:6527–6532.
26. **Kageyama T, et al.** 2004. Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to norovirus in Japan. J. Clin. Microbiol. **42**:2988–2995.
27. **Kapikian AZ, et al.** 1972. Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. J. Virol. **10**:1075–1081.
28. **Kiessling LL, Young T, Gruber TD, Mortell KH.** 2008. Glycoscience 12. **4**:2483–2523.
29. **Levin SA, Dushoff J, Plotkin JB.** 2004. Evolution and persistence of influenza A and other diseases. Math. Biosci. **188**:17–28.
30. **Lindesmith LC, Donaldson EF, Baric RS.** 2011. Norovirus GII.4 strain antigenic variation. J. Virol. **85**:231–242.
31. **Lindesmith LC, et al.** 2008. Mechanisms of GII.4 norovirus persistence in human populations. PLoS Med. **5**:e31.
32. **Marionneau S, et al.** 2001. ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. Biochimie **83**:565–573.
33. **Marotte K, et al.** 2007. X-ray structures and thermodynamics of the interaction of PA-IIL from Pseudomonas aeruginosa with disaccharide derivatives. Chem. Med. Chem. **2**:1328–1338.
34. **Mayer M, Meyer B.** 2001. Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. J. Am. Chem. Soc. **123**:6108–6117.
35. **McCoy AJ, et al.** 2007. Phaser crystallographic software. J. Appl. Crystallogr. **40**:658–674.
36. **Meinecke R, Meyer B.** 2001. Determination of the binding specificity of an integral membrane protein by saturation transfer difference NMR: RGD peptide ligands binding to integrin alphaIIbbeta3. J. Med. Chem. **44**:3059–3065.
37. **Meyer B, Peters T.** 2003. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. Angew Chem. Int. ed. Engl. **42**:864–890.
38. **Mitchell EP, et al.** 2005. High affinity fucose binding of Pseudomonas aeruginosa lectin PA-IIL: 1.0 A resolution crystal structure of the complex combined with thermodynamics and computational chemistry approaches. Proteins **58**:735–746.
39. **Morris GM, et al.** 1998. Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. J. Comput. Chem. **19**:1639–1662.
40. **Neffe AT, Bilang M, Meyer B.** 2006. Synthesis and optimization of peptidomimetics as HIV entry inhibitors against the receptor protein CD4 using STD NMR and ligand docking. Org. Biomol. Chem. **4**:3259–3267.
41. **Otwinowski Z, Minor W.** 1997. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. **276**:307–326.
42. **Penniston KL, Nakada SY, Holmes RP, Assimos DG.** 2008. Quantitative assessment of citric acid in lemon juice, lime juice, and commercially-available fruit juice products. J. Endourol. **22**:567–570.
43. **Prasad BV, et al.** 1999. X-ray crystallographic structure of the Norwalk virus capsid. Science **286**:287–290.
44. **Rademacher C, et al.** 2011. Targeting norovirus infection-multivalent entry inhibitor design based on NMR experiments. Chemistry **17**: 7442–7453.
45. **Rademacher C, Krishna NR, Palcic M, Parra F, Peters T.** 2008. NMR experiments reveal the molecular basis of receptor recognition by a calicivirus. J. Am. Chem. Soc. **130**:3669–3675.