SUPPLEMENTARY INFORMATION APPENDIX

Structural basis for murine norovirus engagement of bile acids and the CD300lf receptor

Christopher A. Nelson[a,*], Craig B. Wilen[a,b,*], Ya-Nan Dai[a,*], Robert C. Orchard[a,c], Arthur S. Kim[a], Roderick A. Stegeman[a], Leon L. Hsieh[a], Thomas J. Smith[d], Herbert W. Virgin[a,e,f,#] & Daved H. Fremont[a,e,g,#]

[a]Department of Pathology & Immunology, Washington University School of Medicine, St. Louis, MO 63110
[b]Present address: Departments of Laboratory Medicine and Immunobiology Yale University New Haven CT 06520
[c]Present address: Department of Immunology, University of Texas Southwestern Medical Center, Dallas, TX 75390
[d]Department of Biochemistry & Molecular Biology, University of Texas Medical Branch, Galveston, TX 77555
[e]Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110
[f]Present address: Vir Biotechnology, San Francisco CA 94158
[g]Department of Biochemistry & Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110

[*]These authors contributed equally.

[#]Correspondence: fremont@wustl.edu or svirgin@vir.bio.

This PDF file includes:

Supplementary Figures S1 to S9
Supplementary Tables S1 and S2
References for SI reference citations

Figure S1

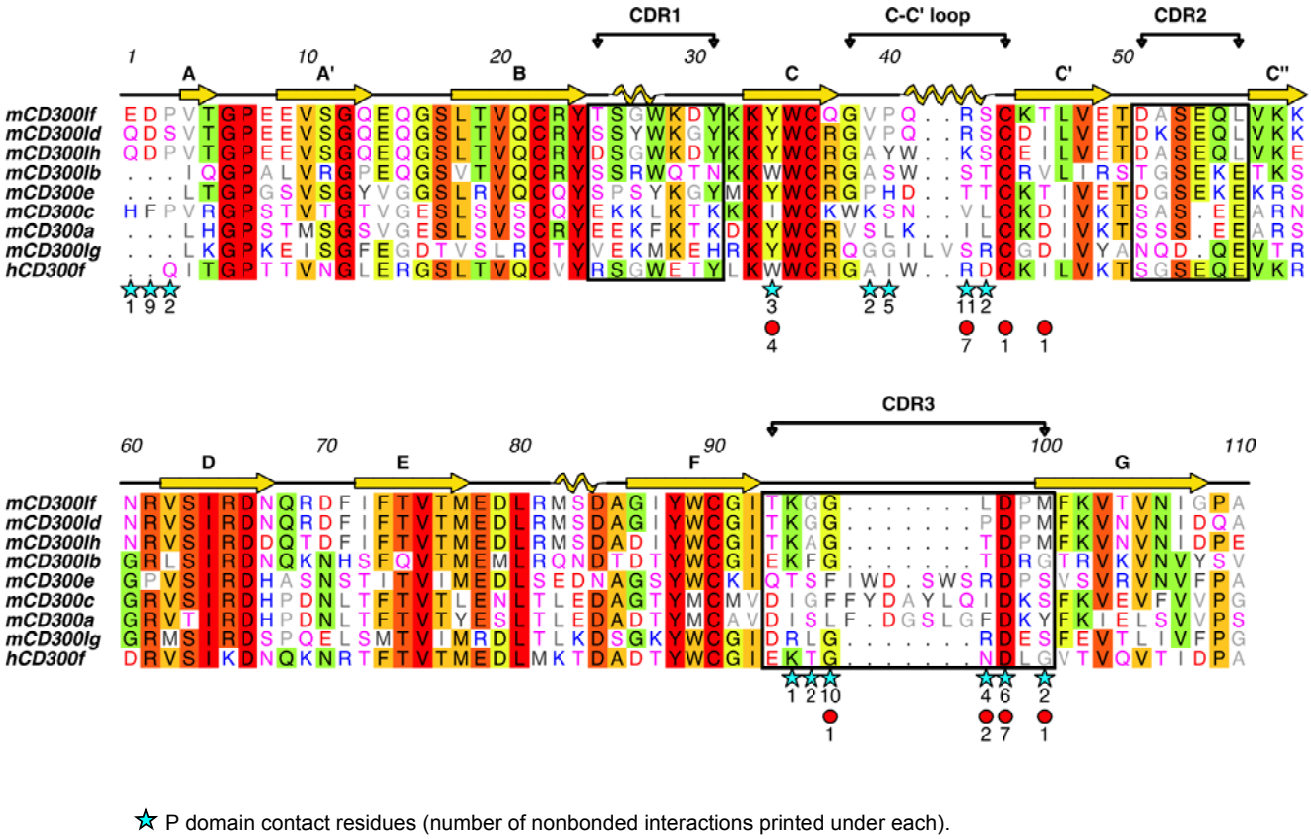
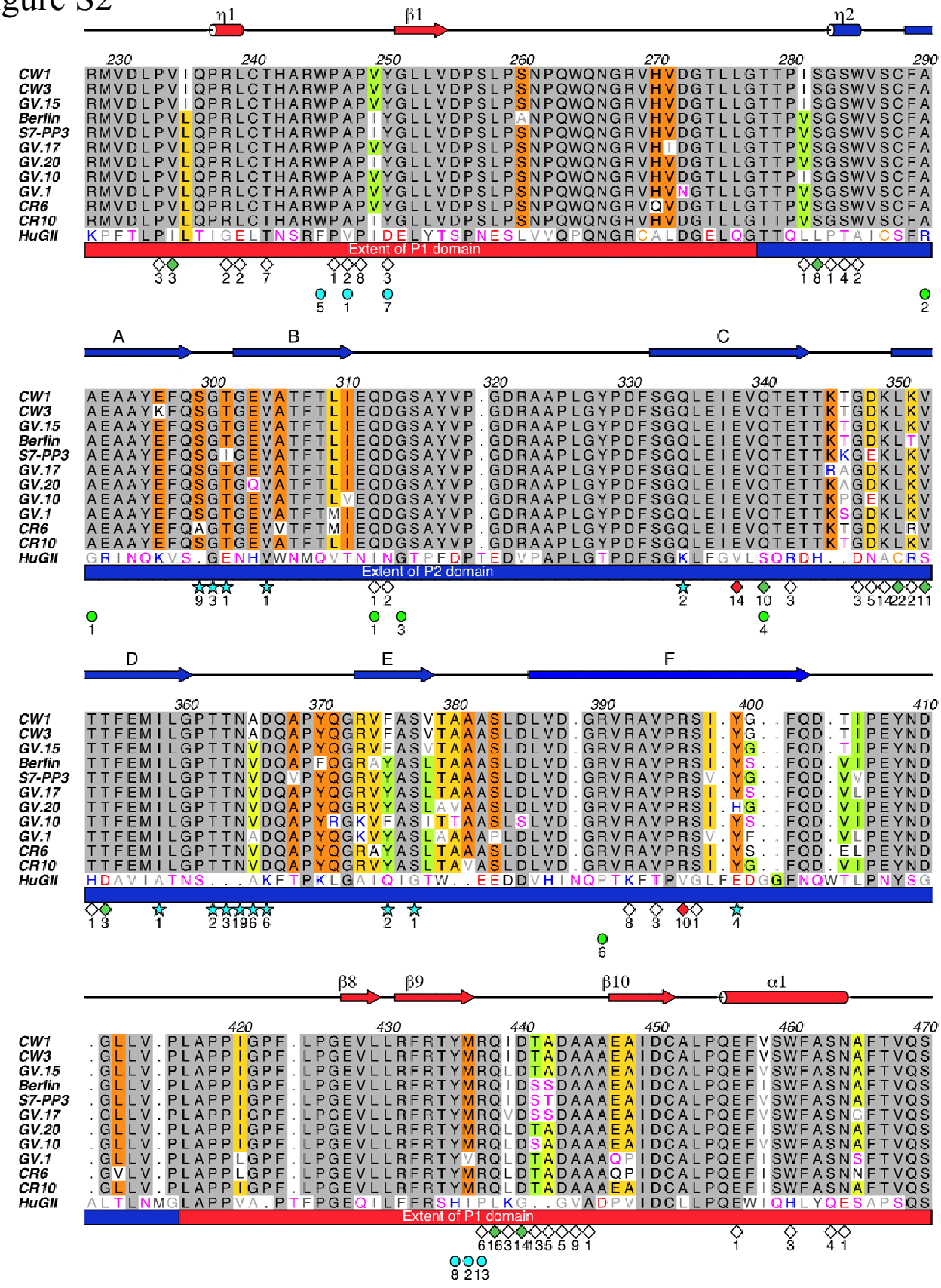


Figure S1. Alignment of CD300 family members. The secondary structure for mCD300lf is shown above the alignment. The complementarity-determining regions (CDR1, CDR2, and CDR3) are marked as defined for the variable chain of an immunoglobulin. The cyan colored stars under the alignment denote positions of contact between mCD300lf and MNoV P domain. The number of nonbonded interactions is given under each contact (star or circle) as determined by HBPLUS using the default cutoff distance of 3.9 Å (1). The sequence conservation is marked by the color of the boxes around the one letter amino acid abbreviations. An "m" in a sequence name stands for mouse, and "h" stands for human. The sequences, obtained from the National Center for Biotechnology Information (NCBI), are as follows: mCD300lf, NCBI reference sequence NP_663609; mCD300ld, reference sequence NP_663412; mCD300lh, reference sequence NP_001095126; mCD300lb Genbank accession AAR27944; mCD300e, reference sequence NP_742047; mCD300c, reference sequence NP_954695; mCD300a, reference sequence NP_739564; mCD300lg, reference sequence, NP_001154183; and hCD300f reference sequence, NP_001276013.

Figure S2



☆ CD300lf contacts (number of nonbonded interactions beneath)
● Contacts with GCDCA from green monomer
● Contacts with cyan monomer
◇ P domain dimer contacts
◆ Hydrogen bond dimer contacts
◆ Salt bridge dimer contacts

Figure S2 continued...

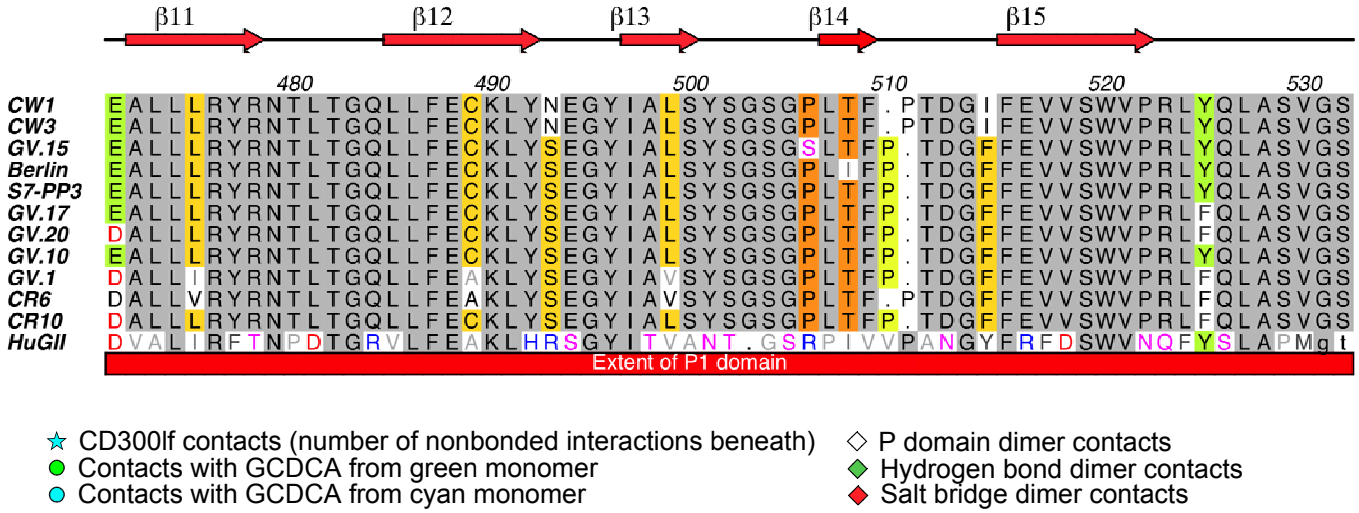


Figure S2. Alignment of norovirus P domain isolates. The secondary structure for MNoV isolate MNV-1.CW3, corresponding to the CD300lf bound x-ray structure in Figure 1, is given above the sequence alignment. The cyan stars mark contacts against CD300lf in the co-complex structure. Diamonds mark the P-domain monomer-monomer contacts at the dimer interface. Green filled diamonds indicate hydrogen bonding positions while red filled diamonds mark salt bridge positions across the dimer interface. The number of nonbonded interactions is given under each contact (diamond, circle, or star) as determined by HBPLUS using the default cutoff distance of 3.9 Å. The sequences were obtained form the National Center for Biotechnology Information (NCBI) using the following Genbank accession numbers: MNV1.CW3, ABJ98944.1; MNV1.CW1, was taken from NCBI reference sequence YP_720002.1; GV.15, ABU55622; Berlin, ABI95835; S7-PP3, BAH04377; GV.17, ABU55598; GV.20, ABU55619; GV.10, ABU55610; GV.1, ABU55586; CR6, AEY83583.1; CR10, ABU55613. HuGII from NCBI reference sequence YP_009237898.

Figure S3

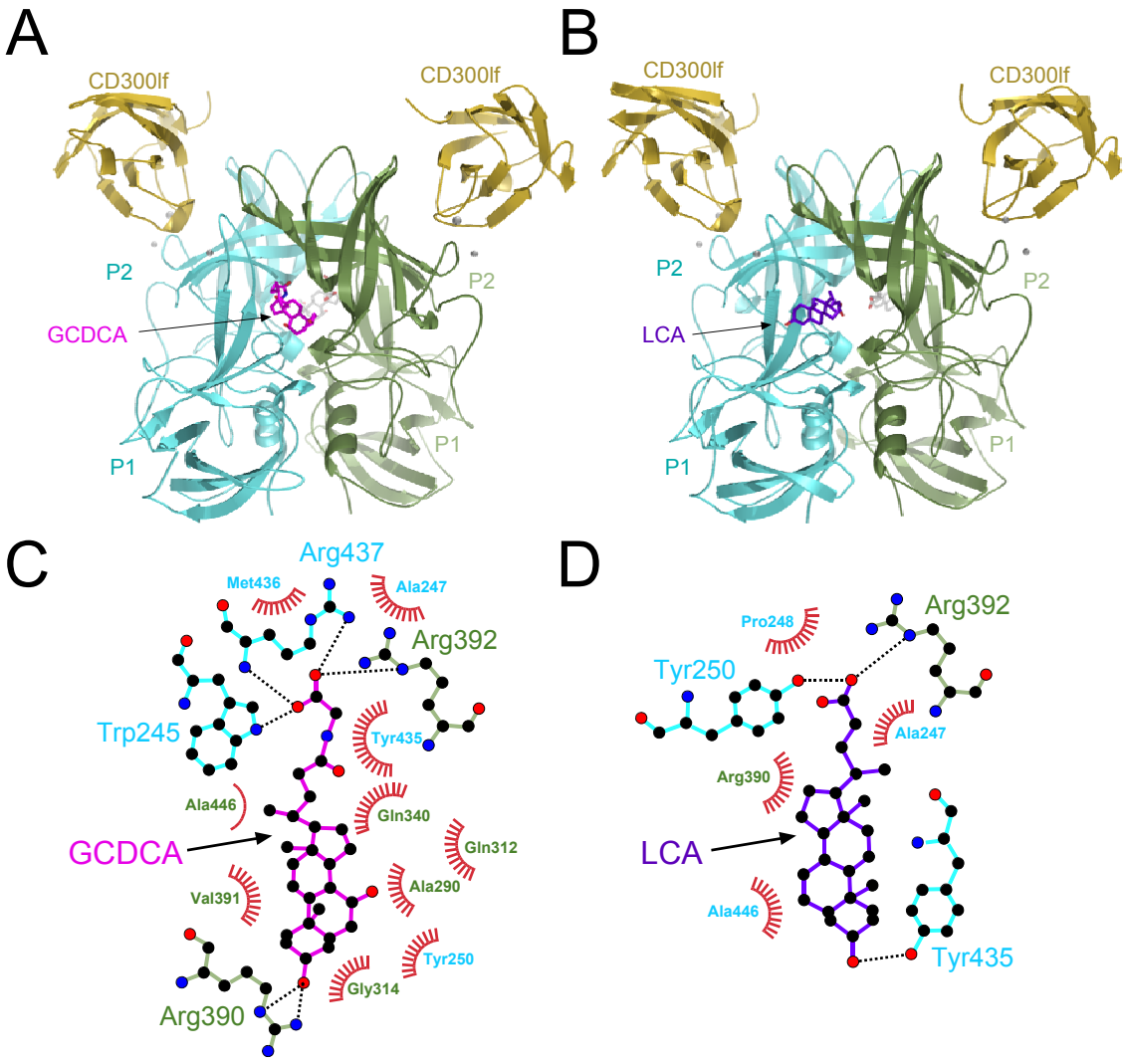
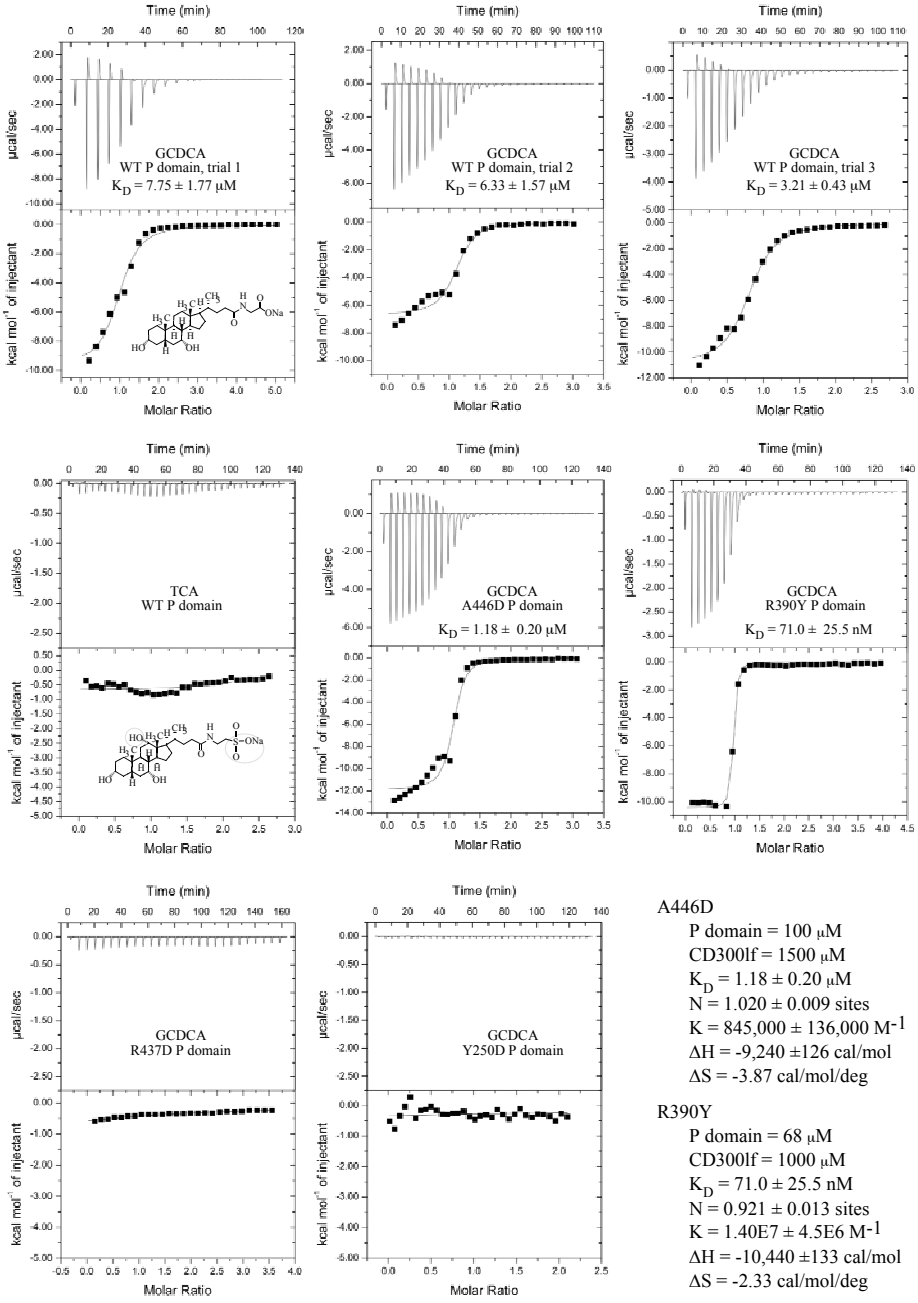


Figure S3. Comparison of GCDCA and LCA binding to the P domain. (A) Ribbon diagram of GCDCA bound to the CD300lf-P domain cocomplex from Fig. 3. (B) Ribbon diagram of LCA bound to the cocomplex. A plot of contacts made by each ligand is given in (C) and (D). The color of the residue name (cyan or green) matches contributing P monomer. Salt bridge interactions are shown as dotted lines. P domain residues that make hydrophobic contacts are shown with red arcs.

Figure S4



A446D
P domain = 100 µM
CD300lf = 1500 µM
K_D = 1.18 ± 0.20 µM
N = 1.020 ± 0.009 sites
K = 845,000 ± 136,000 M^{-1}
ΔH = -9,240 ±126 cal/mol
ΔS = -3.87 cal/mol/deg

R390Y
P domain = 68 µM
CD300lf = 1000 µM
K_D = 71.0 ± 25.5 nM
N = 0.921 ± 0.013 sites
K = 1.40E7 ± 4.5E6 M^{-1}
ΔH = -10,440 ±133 cal/mol
ΔS = -2.33 cal/mol/deg

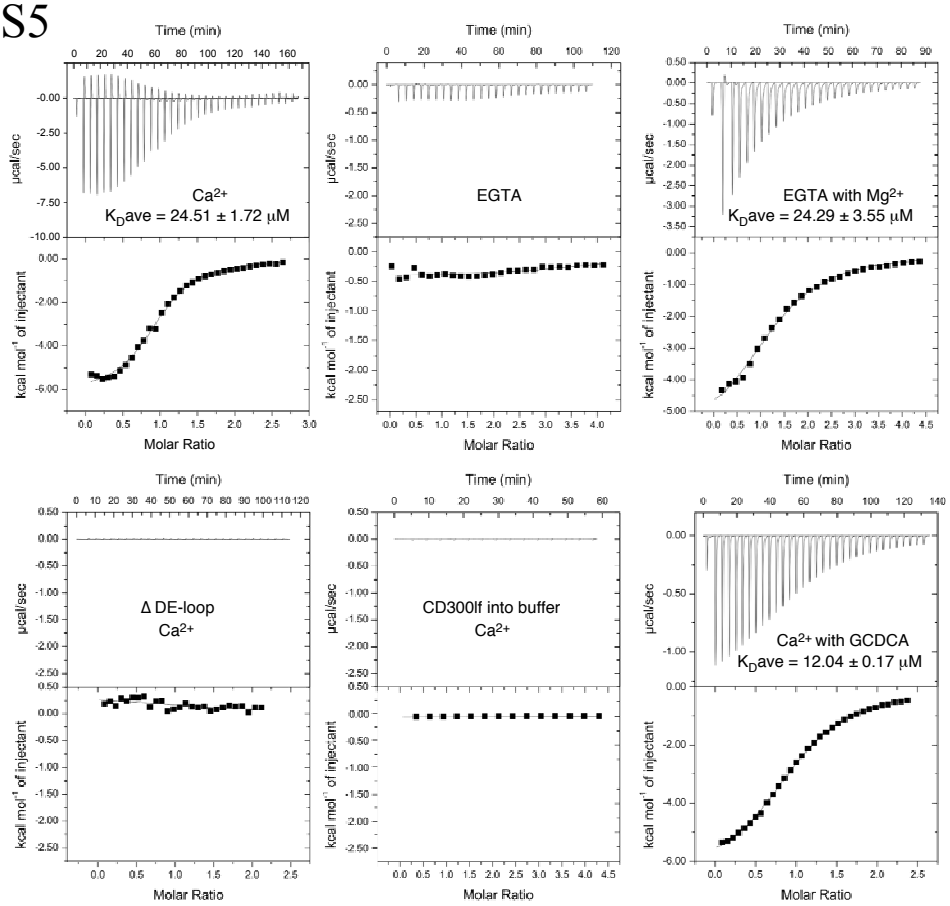
GCDCA, trial 1 (upper row)
P domain = 100 µM
CD300lf = 2599 µM
K_D = 7.75 ± 1.77 µM
N = 0.911 ± 0.011 sites
K = 129,000 ± 28,000 M^{-1}
ΔH = -9706 ± 414 cal/mol
ΔS = -9.18 cal/mol/deg

GCDCA, trial 2
P domain = 100 µM
CD300lf = 1500 µM
K_D = 6.33 ± 1.57 µM
N = 0.913 ± 0.029 sites
K = 158,000 ± 37,100 M^{-1}
ΔH = -10,360 ± 324 cal/mol
ΔS = -10.9 cal/mol/deg

GCDCA, trial 3
P domain = 100 µM
CD300lf = 1000 µM
K_D = 3.21 ± 0.43 µM
N = 0.803 ± 0.013 sites
K = 311,000 ± 41,300 M^{-1}
ΔH = -10,950 ±104 cal/mol
ΔS = -11.60 cal/mol/deg

Figure S4. Binding affinity of GCDCA for P domain, and P domain variants as measured by isothermal titration calorimetry. The P domain proteins were dialyzed extensively against buffer containing 30 mM Hepes, pH 7.4, 150 mM NaCl and 1 mM $CaCl_2$. The top row shows injections of GCDCA in the same buffer into the reaction well containing wild type CW3 P domain. The first panel includes a chemical diagram of the sodium salt of GCDCA. The binding affinity was reproducible yielding as the average from 3 trials a K_D of ~ 5.76 ± 1.26 µM. The first panel of the second row shows injection into P domain of TCA, a closely related bile salt that does not enhance MNoV infection. No binding was observed and we conclude the binding signal depends on the exact chemical formula of the bile acid. A chemical diagram of the sodium salt of TCA is included for comparison with GCDCA. The remaining panels show injections of GCDCA into variants of the MNoV P domain. The A446D variant binds similar to wild type. The R390Y binds with ~ 50 fold higher affinity. The remaining two variants, R437D, Y250D show very little if any binding.

Figure S5



Ca^{2+}, trial 1 (shown upper left)	Ca^{2+}, trial 2	
P domain = 294 µM	P domain = 53 µM	
CD300lf = 3964 µM	CD300lf = 2050 µM	
K_D = 27.10 ± 2.54 µM	K_D = 21.93 ± 0.89 µM	$K_D$ave = 24.51 ± 1.72 µM
N = 0.979 ± 0.012 sites	N = 1.070 ± 0.034 sites	
K = 36900 ± 3420 M^{-1}	K = 45600 ± 4020 M^{-1}	
ΔH = -6184 ±104 cal/mol	ΔH = -6714 ± 271.0 cal/mol	
ΔS = 0.291 cal/mol/deg	ΔS = -1.19 cal/mol/deg	
EGTA with Mg^{+2}, trial 1 (upper right)	EGTA with Mg^{+2}, trial 2	
P domain = 96 µM	P domain = 224 µM	
CD300lf = 2000 µM	CD300lf = 2643 µM	
K_D = 36.90 ± 6.51 µM	K_D = 11.68 µM	$K_D$ave = 24.29 ± 3.55 µM
N = 1.320 ± 0.040 sites	N = 1.690 ± 0.064 sites	
K = 27100 ± 2790 M^{-1}	K = 85600 ± 4100 M^{-1}	
ΔH = -5986 ± 247 cal/mol	ΔH = -4343 ±222 cal/mol	
ΔS = 0.208 cal/mol/deg	ΔS = 8.01 cal/mol/deg	
Ca^{2+} with GCDCA, trial 1 (lower left)	Ca^{2+} with GCDCA, trial 2	
P domain = 75 µM	P domain = 53 µM	
CD300lf = 900 µM	CD300lf = 1720 µM	
K_D = 14.27 ± 0.21 µM	K_D = 9.80 ± 0.17 µM	$K_D$ave = 12.04 ± 0.17 µM
N = 1.01 ± 0.01 sites	N = 1.17 ± 0.04 sites	
K = 70100 ± 2010 M^{-1}	K = 102000 ± 1760 M^{-1}	
ΔH = -6574 ± 55.16 cal/mol	ΔH = -6590 ± 270 cal/mol	
ΔS = 0.123 cal/mol/deg	ΔS = 1.48 cal/mol/deg	

Figure S5. Binding affinity of CD300lf for the P domain measured by isothermal titration calorimetry. The upper left panel shows the titration of CD300lf into P domain. The K_D = 24.51 ± 1.72 µM value represents the average of two independent trials. The Ca^{2+} indicates the CD300lf and P domain proteins were dialyzed extensively against buffer containing 30 mM Hepes, pH 7.4, 150 mM NaCl with 1mM $CaCl_2$. For those trials where binding occured, the protein concentrations used, and the parameters obtained are give below the data plots. Dialysis of the proteins in buffer containing 0.1M EGTA without added $CaCl_2$ resulted in no binding signal (middle upper panel). Binding was observed in buffer containing 0.1M EGTA with 1 mM $MgCl_2$ (upper left panel), being nearly equivalent to that observed in the original Ca^{2+} buffer. The DE-loop panel shows CD300lf binding was undetectable to a P domain mutant missing part of the CD300lf binding site (residues Asn364, Ala365, and Asp366). The DE-loop variant does not bind Fc-CD300lf by BLI (Fig 6). The lower middle panel shows no binding was observed upon injection of CD300lf into buffer alone (no P domain control). Although it was not possible to calculate a molar ratio (without P domain) the data was plotted to match other titrations for comparison. In the lower right panel, addition of 250 µM GCDCA to the calcium buffer gave a small but reproducible increase in CD300lf binding affinity, about 2 fold, yielding a K_D 12.1 ± 0.17 µM value. These data indicate that divalent cation(s) are required for CD300lf binding. Binding parameters such as the number of binding sites (n), the binding constant (K, M^{-1}), and the binding enthalpy (ΔH, kcal/mol of bound ligand), were determined by fitting the experimental binding isotherms.

Figure S6

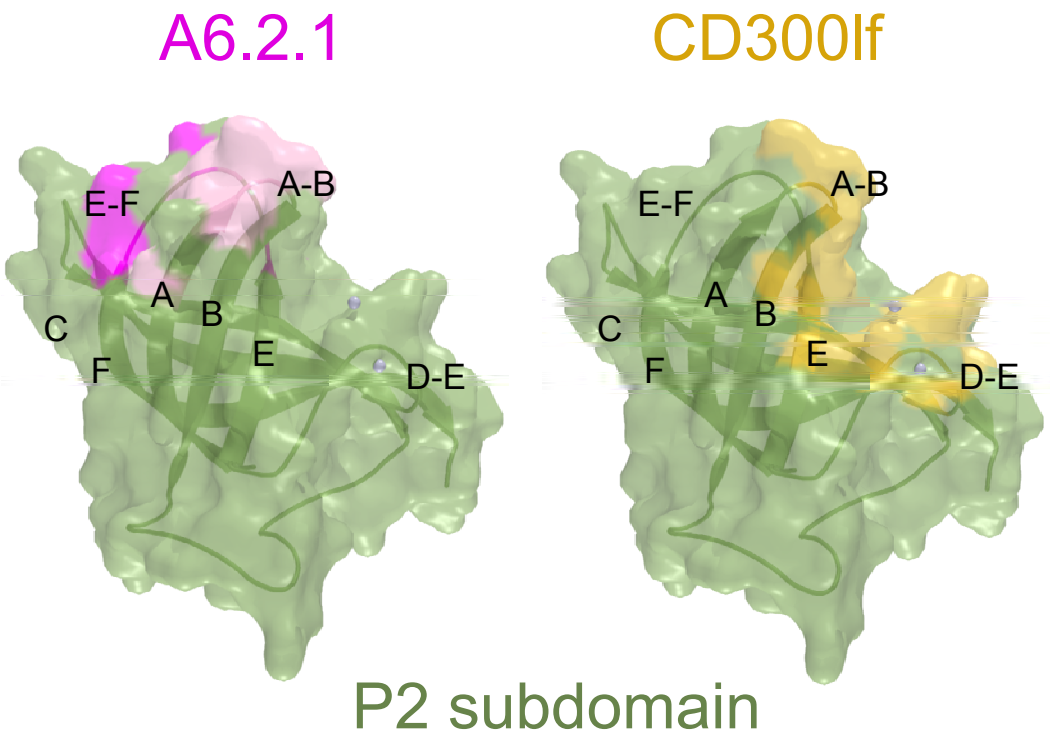


Figure S6. Comparison of predicted A6.2.1 contact site with the CD300lf contact site on the P domain. On the left is the predicted footprint of the A6.2.1 antibody. The dark magenta represents the four positions reported for escape mutants (2). The pink area shows the surface contributed by residues implicated in A6.2.1 binding by site directed mutagenesis studies (3-5). On the left in yellow are the P domain contacts within 4.5 Å of CD300lf determined from the co-complex structure (Fig. 1). The P domain surface is in green with metal ions in silver.

Figure S7

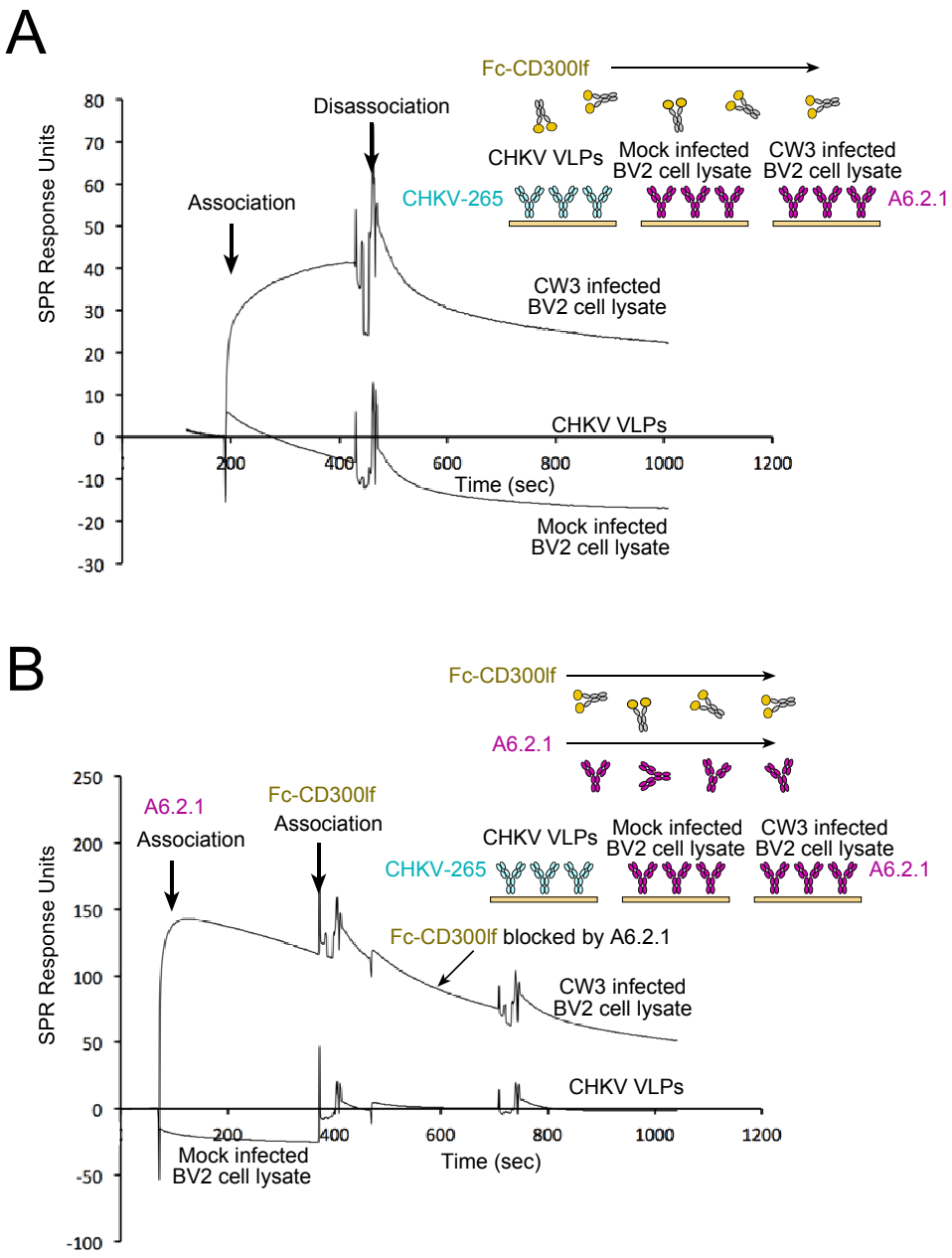


Figure S7. Fc-CD300lf binding to P domain purified from infected cell lysate. To capture virus and virus fragments, the P domain specific antibody A6.2.1 was immobilized on a CM5 chip using standard NHS-EDC chemistry. Then MNV1.CW3-infected cell lysate was passed over the A6.2.1 coated surface (in the last flow cell) and allowed to bind. For a negative control, mock-infected cell lysate was passed over the A6.2.1 coated surface (middle flow cell). For a reference, the Chikungunya virus specific antibody CHKV-265 was immobilized onto the chip and used to capture purified Chikungunya virus-like particles (VLPs, first flow cell). (A) When Fc-CD300lf was passed over all three prepared surfaces, very little binding occurred on the Chikungunya or the mock lysate coated surfaces but good binding was recorded in the last flow cell containing the CW3 virus lysate coated surface. The Chikungunya VLP reference signal was subtracted from all three curves. Arrows mark the start and stop of the Fc-CD300lf injection. (B) To check for specificity, the experiment was repeated but this time the antibody A6.2.1 was used to cover any P domain protein before the Fc-CD300lf was passed over the chip. A6.2.1 did not bind the Chikungunya VLP or mock lysate coated surface, although it did bind strongly to the CW3 lysate coated surface. Further, pretreatment with A6.2.1 significantly reduced binding of Fc-CD300lf to the CW3 lysate, demonstrating that the Fc-CD300lf binding signal is specific for the P domain.

Figure S8

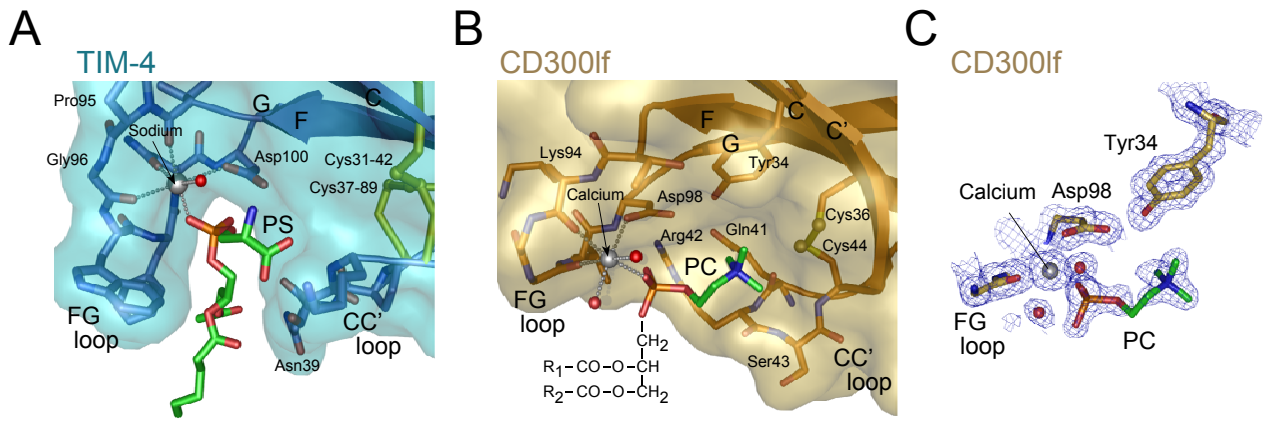


Figure S8. TIM-4 and CD300lf bind lipid headgroups by a similar mechanism. (A) Surface representation of the TIM-4 IgV domain showing the bound ligand phosphatidylserine (PS). The coordinating sodium ion is shown in silver with the bonding as gray dots. One water coordinated to the metal is shown as a red sphere. (B) The CD300lf binding pocket with a phosphocholine (PC) headgroup bound. (C) Electron density map around the PC ligand (in blue mesh, contoured at 1σ).

Figure S9

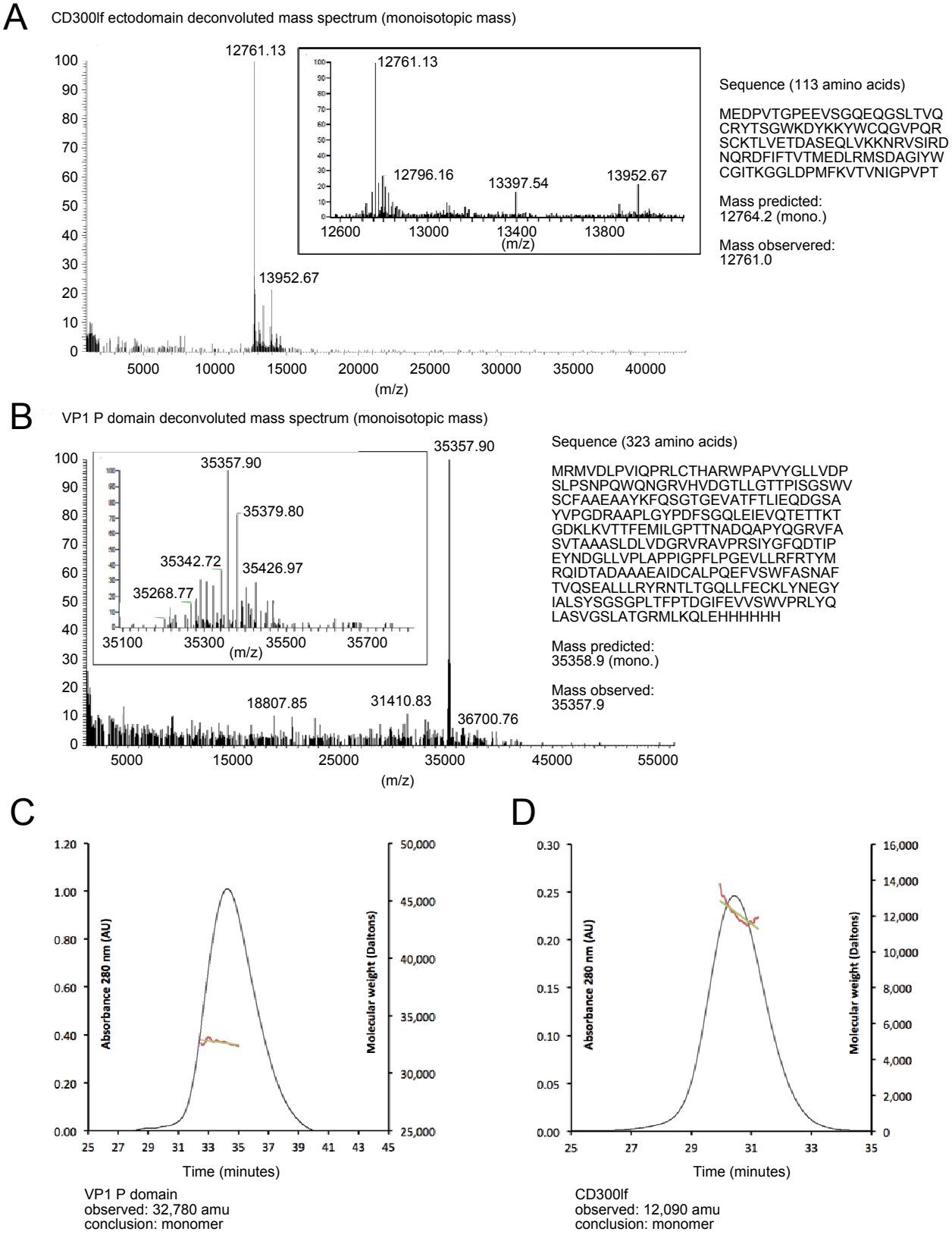


Figure S9. The CD300lf and P domain proteins used in this study are monomeric. (A) The deconvoluted mass spectrum for the CD300lf ectodomain protein determined by electrospray mass spectrometry. On the right is the amino acid sequence of the expression construct showing good agreement between the predicted and observed mass values. (B) The mass spectrometry data and sequence for the VP1 P domain protein. Both the purified P domain (C) and the CD300lf ectodomain (D) assay as monomer by size exclusion chromatrography multi-angle light scatter (SEC-MALS) analysis.

Table S1. Data Collection and Refinement Statistics

	CD300lf / P domain	CD300lf / PC
PDB ID code	6C6Q	6C74
Unit-cell, Å	48.71, 135.24, 139.51	32.64, 40.06, 68.29
Space group	$P2_12_12_1$	$P2_12_12_1$
[a]Resolution range, Å	62.00- 2.00 (2.07 - 2.00)	34.15 - 1.35 (1.41 - 1.35)
Total reflections	377874 (17218)	117261 (5469)
Unique reflections	57592 (3530)	19612 (1620)
Average multiplicity	6.6 (4.8)	6.0 (3.2)
Mean I/σ(I)	8.56 (0.79)	11.21 (0.54)
Completeness, %	91 (57)	98 (83)
Rmerge	0.1781 (1.7)	0.1039 (1.802)
Rmeas	0.1937 (1.913)	0.1139 (2.169)
CC1/2	0.994 (0.439)	0.998 (0.421)
CC*	0.998 (0.781)	0.999 (0.77)
Refinement		
Rwork	0.1847 (0.2899)	0.1677 (0.3581)
[b]Rfree	0.2253 (0.3482)	0.1919 (0.3306)
CC(work)	0.956 (0.727)	0.971 (0.428)
CC(free)	0.932 (0.600)	0.963 (0.405)
Number of Atoms		
non-hydrogen atoms	7267	1028
macromolecules	6443	877
ligands	96	13
Protein residues	828	111
Solvent molecules	728	138
RMS(bonds)	0.004	0.012
RMS(angles)	0.60	1.34
Ramachandran, favored, %	96.59	97.25
Ramachandran, allowed, %	3.41	2.75
Ramachandran, outliers, %	0	0.00
Rotamer outliers, %	0.71	0.00
Clashscore	0.62	2.84
B-factor Model		
Average B factor, Å^2	30.58	16.76
macromolecules	29.99	15.19
ligands	41.60	22.76
solvent	34.40	26.19

[a]Values in parentheses refer to the highest resolution shell. [b]Rfree = free R factor based on random 5% of all data. Diffraction source was ALS BL4.2.2 using detector RDI CMOS_8M. Data processing, scaling statistics, and refinement statistics are described in the Experimental Procedures.

Table S2. Data Collection and Refinement Statistics for Bile Acid Structures

	CD300lf / P domain GCDCA	CD300lf / P domain LCA
PDB ID code	**6E47**	**6E48**
Unit-cell, Å	49.11, 134.82, 139.59	101.09 152.89 70.04
Space group	$P2_12_12_1$	$P2_12_12$
[a]Resolution range, Å	60.70- 1.95 (2.02 - 1.95)	63.67 - 1.80 (1.87 - 1.80)
Total reflections	387544 (15173)	188776 (11228)
Unique reflections	67388 (1306)	95932 (6238)
Average multiplicity	5.8 (2.6)	2.0 (1.8)
Mean I/σ(I)	6.48 (0.14)	5.92 (0.65)
Completeness, %	87.04 (19.34)	95 (63)
Rmerge	0.2444 (5.4)	0.0688 (0.803)
Rmeas	0.2694 (6.8)	0.0972 (1.135)
CC1/2	0.991 (0.046)	0.996 (0.324)
CC*	0.998 (0.297)	0.999 (0.699)
Refinement		
Rwork	0.2085 (0.3286)	0.1743 (0.3354)
[b]Rfree	0.2570 (0.3262)	0.2039 (0.3711)
CC(work)	0.953 (0.501)	0.963 (0.558)
CC(free)	0.935 (0.417)	0.955 (0.364)
Number of Atoms		
non-hydrogen atoms	7041	7622
macromolecules	6394	426
ligands	158	82
Protein residues	825	830
Solvent molecules	489	138
RMS(bonds)	0.002	0.008
RMS(angles)	0.62	0.97
Ramachandran, favored, %	96.21	96.23
Ramachandran, allowed, %	3.79	3.77
Ramachandran, outliers, %	0	0.00
Rotamer outliers, %	0.72	0.86
Clashscore	0.54	1.32
B-factor Model		
Average B factor, Å^2	34.65	22.50
macromolecules	34.17	20.84
ligands	45.01	24.66
solvent	37.52	31.88

[a]Values in parentheses refer to the highest resolution shell. [b]Rfree = free R factor based on random 5% of all data. Diffraction source was ALS BL4.2.2 using detector RDI CMOS_8M. Data processing, scaling statistics, and refinement statistics are described in the Experimental Procedures.

References

1. McDonald IK, Thornton JM (1994). Satisfying hydrogen bonding potential in proteins. J. Mol. Biol. 1994;238:777-793.

2. Kolawole AO, Li M, Xia C, Fischer AE, Giacobbi NS, Rippinger CM, Proescher JB, Wu SK, Bessling SL, Gamez M, Yu C, Zhang R, Mehoke TS, Pipas JM, Wolfe JT, Lin JS, Feldman AB, Smith TJ, Wobus CE. 2014. Flexibility in surface-exposed loops in a virus capsid mediates escape from antibody neutralization. J Virol 88:4543–4557. doi:10.1128/JVI.03685-13.

3. Kolawole AO, Smith HQ, Svoboda SA, Lewis MS, Sherman MB, Lynch GC, Pettitt BM, Smith TJ, Wobus CE. Norovirus Escape from Broadly Neutralizing Antibodies Is Limited to Allostery-Like Mechanisms. mSphere. 2017;2(5). doi:10.1128/mSphere.00334-17. PubMed PMID: 29062895; PMCID: PMC5646240.

4. Taube S, Rubin JR, Katpally U, Smith TJ, Kendall A, Stuckey JA, Wobus CE. 2010. High-resolution X-ray structure and functional analysis of the murine norovirus 1 capsid protein protruding domain. J Virol84:5695–5705. doi:10.1128/JVI.00316-10. PubMed PMID: 20335262; PMCID: PMC2876589.

5. Lochridge VP, Hardy ME. A single-amino-acid substitution in the P2 domain of VP1 of murine norovirus is sufficient for escape from antibody neutralization. J Virol. 2007;81(22):12316-22. doi: 10.1128/JVI.01254-07. PubMed PMID: 17804495; PMCID: PMC2168968.