Short Article

Structure of the divergent human astrovirus MLB capsid spike

Graphical abstract
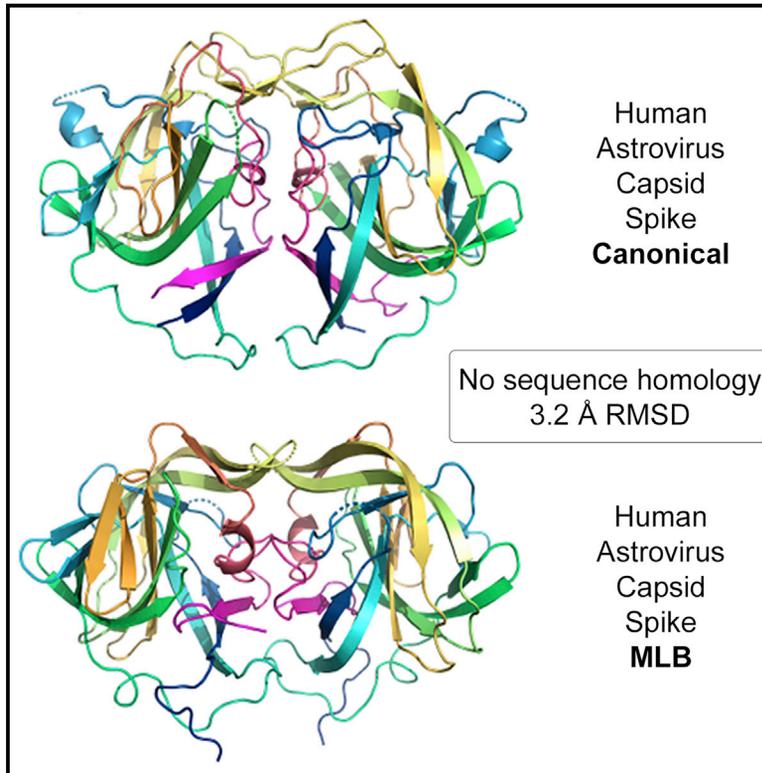
Authors

Kevin Delgado-Cunningham,
Tomás López, Firas Khatib,
Carlos F. Arias, Rebecca M. DuBois

Correspondence

rmdubois@ucsc.edu

In brief

Delgado-Cunningham et al. report the structure of the capsid spike from human astrovirus strain MLB1. The capsid spike forms protrusions on the surface of the virus and may have a role in attaching to human cells. The MLB1 spike structure has many differences compared with classical human astrovirus spike structures.

Highlights

- Crystal structure of the human astrovirus MLB capsid spike domain is reported

- Classical and MLB human astroviruses have structurally divergent capsid spikes

- Competition data support the existence of multiple human astrovirus receptors

- The MLB spike was a challenging target for structure prediction

CellPress

Structure

Short Article

Structure of the divergent human astrovirus MLB capsid spike

Kevin Delgado-Cunningham,[1] Tomás López,[2] Firas Khatib,[3] Carlos F. Arias,[2] and Rebecca M. DuBois[1,4,*]
[1]Department of Biomolecular Engineering, University of California, Santa Cruz, Santa Cruz, CA 95064, USA
[2]Departamento de Genética del Desarrollo y Fisiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, 62210, Mexico
[3]Department of Computer and Information Science, University of Massachusetts Dartmouth, Dartmouth, MA 02747, USA
[4]Lead contact
*Correspondence: rmdubois@ucsc.edu
https://doi.org/10.1016/j.str.2022.10.010

SUMMARY

Despite their worldwide prevalence and association with human disease, the molecular bases of human astrovirus (HAstV) infection and evolution remain poorly characterized. Here, we report the structure of the capsid protein spike of the divergent HAstV MLB clade (HAstV MLB). While the structure shares a similar folding topology with that of classical-clade HAstV spikes, it is otherwise strikingly different. We find no evidence of a conserved receptor-binding site between the MLB and classical HAstV spikes, suggesting that MLB and classical HAstVs utilize different receptors for host-cell attachment. We provide evidence for this hypothesis using a novel HAstV infection competition assay. Comparisons of the HAstV MLB spike structure with structures predicted from its sequence reveal poor matches, but template-based predictions were surprisingly accurate relative to machine-learning-based predictions. Our data provide a foundation for understanding the mechanisms of infection by diverse HAstVs and can support structure determination in similarly unstudied systems.

INTRODUCTION

Astroviruses are a family of nonenveloped, positive-sense RNA viruses that infect a variety of avian and mammalian species. Three astrovirus clades—classical, VA, and MLB—are known to infect humans (Figure 1A).[1] Human astroviruses (HAstVs) from all three clades are believed to replicate in the intestinal tract and may cause both gastrointestinal and extra-gastrointestinal disease.[2,3] Classical HAstVs infect the intestinal epithelium[4] and are a well-established cause of diarrhea, particularly in pediatric patients.[3,5] Infection by HAstV VA1 may cause diarrhea[1,6–8] and severe inflammation in the central nervous systems of immunocompromised patients.[9,10] HAstV MLBs (including MLB1, MLB2, and MLB3) are found in stool samples worldwide,[11–14] and seroprevalence to HAstV MLB capsid is nearly ubiquitous in humans.[15] However, case-control experiments testing the link between HAstV MLB infection and gastroenteritis have yielded conflicting results. HAstV MLB and/or HAstV MLB1 infection was not associated with gastroenteritis in two studies,[12,14] whereas HAstV MLB1 infection was associated with gastroenteritis in another study.[13] Several case studies link HAstV MLB to viremia[16–23] and disease in the CNS.[18,21,24] Overall, a conclusive link of HAstV MLB to disease has yet to be established.

As HAstV MLBs are found in stool samples and can be cultured in human intestinal enteroids,[25] they are believed to initially infect the intestinal epithelium *in vivo*. Diverse astroviruses have been shown to infect enterocytes, including classical,[4] VA1,[25] ovine,[26] murine,[27,28] bovine,[29] and turkey astroviruses.[30–32] Interestingly, all 8 classical HAstV serotypes (HAstV 1–8) and HAstV VA1 can be cultured in the human colonic cancer cell line Caco-2,[33–35] but the only published effort to grow HAstV MLB1 in Caco-2 cells was unsuccessful.[36] Instead, HAstV MLBs could be propagated in two human nonintestinal cell lines, Huh-7 and A549; however, these cells are only partially permissive, with infection observed in ~10% of cells.[36] The molecular drivers of these observed differences in cell tropism are unknown, as are their contributions in human disease.

The HAstV capsid is a T = 3 icosahedron made up of 180 copies of the capsid protein.[37] The capsid basic and core domains encase the viral genome, while spike domain dimers project outward from the surface[38–40] and have been shown to mediate cell attachment.[41,42] Classical HAstVs then enter Caco-2 cells (and potentially host cells *in vivo*) through clathrin-mediated endocytosis.[43] Before viral release from host cells, classical HAstV capsid protein undergoes intracellular proteolysis by host caspases to remove a C-terminal acidic domain.[44,45] Upon viral release, further digestion by trypsin *in vitro* yields an infectious particle containing 30 well-ordered spike dimers on the icosahedral 2-fold axes.[37–39] Extracellular proteolysis using trypsin is required for propagation of classical but not VA or MLB HAstVs in cell culture.[34,36,46] Mature HAstV VA1 is instead

Structure
Short Article
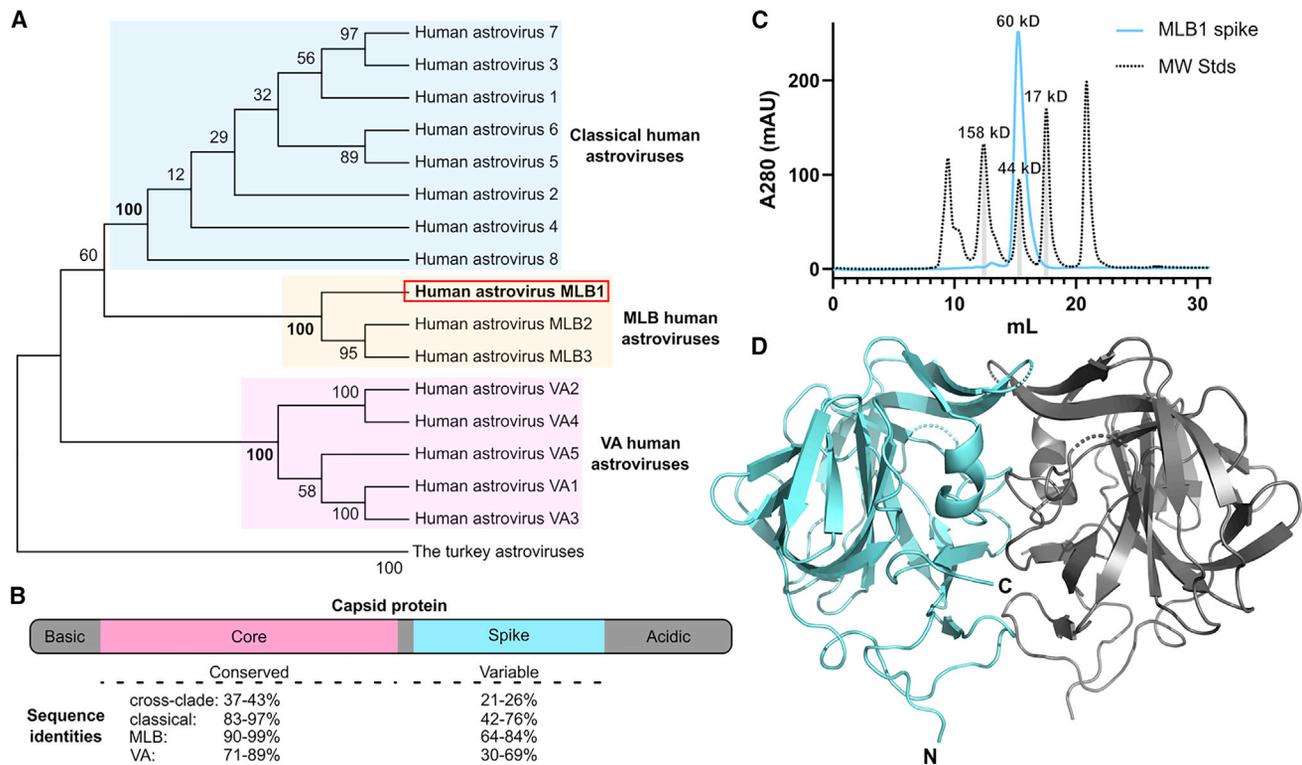
Figure 1. Delineation and structure of the HAstV MLB spike domain

(A) The human astroviruses. Complete capsid protein sequences were aligned using the MUSCLE algorithm. Evolutionary analyses in MEGA-X using the maximum likelihood method and JTT matrix-based model yielded the cladogram shown (log likelihood −21,535.53). Bootstrap values are shown next to the branches and were computed using 1,000 replicates. The tree was rooted using the turkey astrovirus 1–3 capsid sequences as an outgroup.

(B) The domain structure of the astrovirus capsid protein. Pairwise sequence identities in the core and spike domains between and within the HAstV clades (classical, MLB, VA) are included. The diagram is drawn so that domains are arranged from N-terminal (basic) to C-terminal (acidic).

(C) Size-exclusion chromatography trace of recombinant HAstV MLB1 spike (cyan line) overlaid with the trace of gel filtration standards (black dotted line).

(D) Crystal structure of the MLB1 spike dimer.

produced through intracellular, caspase-independent cleavage by one or more unknown proteases.[47] The proteolytic maturation of MLB HAstVs is currently uncharacterized.

While no astrovirus receptors have been identified to date, structural studies of classical HAstV spikes have yielded sites on the spike surface that may function in receptor binding.[40–42] Of these, the P site is of particular interest. This patch of conserved residues is located at the top of the spike, between the epitopes of antibodies known to block viral attachment.[40,41] The hydrophilicity of the P site, including several hydrogen-bond donors and acceptors, supported the hypothesis that HAstV may attach to host cells by binding a carbohydrate.[40] While heparin, dextran sulfate, heparan sulfate, and chondroitinase treatment of host cells partially inhibit HAstV infection,[40,48] if HAstVs use or require carbohydrates for infection *in vivo* is unknown. Crucially, our current understanding of astrovirus cell entry has been built using primarily classical HAstVs. The molecular basis for entry by HAstV VA and MLB remains wholly unstudied.

Here, we delineated the HAstV MLB capsid spike domain and solved its structure as a first step toward understanding the molecular basis for cell entry by HAstV MLB. While the HAstV MLB spike shares a folding topology with the spike domain of classical HAstV, it is otherwise strikingly different. None of the putative receptor-binding sites from classical HAstV spikes were

observed on the surface of the MLB spike, supporting the hypothesis that the MLB and classical HAstVs require different host factors for cell entry. To test this, we developed a novel *in vitro* HAstV infection competition assay and provide evidence that HAstV MLB utilizes a different receptor for cell infection than classical and VA HAstVs. Our data lay a foundation for elucidating the cell entry mechanisms used by divergent HAstVs.

RESULTS

Delineation of the MLB spike domain

To delineate the boundaries of the HAstV MLB capsid spike domain, we evaluated the capsid protein sequences of classical, MLB, and VA1 HAstVs (Figure 1A). Sufficient identity between the HAstV capsid core domain sequences (37%–43%), along with the classical HAstV core domain structures, allowed the delineation of the core domain boundaries (Figure 1B).[38,39] However, the HAstV MLB spike domain boundaries could not be precisely mapped by comparison to classical HAstV spike sequences and structures due to low sequence identity (22%–26%) (Figure 1B).[38,40,42,49] Thus, we tested several recombinant HAstV MLB1 capsid spike domain constructs and identified a construct containing residues 399–653 of the HAstV MLB1 capsid protein (30 kD) that could be expressed in Sf9 insect cells

Structure
Short Article

Table 1. Crystallographic statistics

Parameter	MLB1 spike[a]	Se-Met MLB1 spike[a]
PDB Code	7UZT	–
Data collection		
Wavelength (Å)	1.00004	0.97942
Space group	P 1 2 1	P 1 2 1
Cell dimensions		
a, b, c (Å)	72.10, 77.11, 75.44	72.15, 77.07, 75.44
α, β, γ (°)	90, 110.71, 90	90, 110.75, 90
Resolution (Å)	47.66–1.86 (1.90–1.86)	47.67–2.53 (2.63–2.53)
R_{merge}	0.082 (0.433)	0.425 (4.975)
I/σI	8.7 (2.7)	14.7 (5.2)
Completeness (%)	98.1 (94.2)	99.8 (98.2)
Redundancy	3.4 (3.3)	27.1 (26.2)
$CC_{1/2}$	0.995 (0.884)	0.995 (0.612)
No. of Selenium sites	–	27
Overall figure of merit	–	0.382
Refinement		
No. of reflections	63,794	–
Resolution (Å)	47.66–1.86	–
R_{work}/R_{free}	0.242/0.260	–
No. of atoms	5,271	–
Protein	5,040	–
Ligands	0	–
Water	231	–
B factors (Å²)	22.50	–
Protein (Å²)	22.34	–
Ligands (Å²)	N/A	–
Water (Å²)	26.08	–
RMSD		
Bond lengths (Å)	0.011	–
Bond angles (°)	1.449	–
Ramachandran statistics		
Favored (%)	96.28	–
Allowed (%)	3.23	–
Outliers (%)	0.48	–

Se-Met MLB1 spike, selenomethionine-substituted MLB1 spike; RMSD: root-mean-square deviation.
[a]Data from one crystal were used to solve each structure. Statistics in parentheses correspond to the highest-resolution shell in each dataset.

and *Escherichia coli* and remains soluble after affinity purification. Analysis of this construct by size-exclusion chromatography revealed an elution volume that is consistent with a homodimer (∼60 kD), as observed for other astrovirus capsid spike domains (Figure 1C).[40,48]

Structure of MLB spike and comparison to the classical HAstV spike

To gain insights into the function of the HAstV MLB1 spike, we used X-ray crystallography to determine its high-resolution

structure. The HAstV MLB1 spike crystallized readily and yielded X-ray diffraction data to a resolution of 1.86 Å (Table 1), but we were unable to solve its structure by molecular replacement with the classical HAstV spike structure or with predicted models (discussed further below). Instead, we generated a selenomethionine-substituted HAstV MLB1 spike and used single-wavelength anomalous diffraction to solve the structure (Table 1). Structure determination revealed three HAstV MLB1 spike chains (A–C) in the asymmetric unit, with pairwise root-mean-square deviations (RMSDs) between 0.18 and 0.34 Å. Each HAstV MLB1 spike chain is comprised of ordered residues 420–645, except for two or three disordered loops (discussed below). Each HAstV MLB1 spike chain forms a homodimer through crystallographic symmetry, consistent with size-exclusion chromatography data (Figure 1D).

The HAstV MLB1 spike homodimer has ∼3,100 Å² buried at the dimer interface (Figure 2A), similar to the classical spike, which has a 3,500–3,800 Å² dimer interface.[38,40,42,49] The structural scaffold of each protomer is anchored by an antiparallel β-barrel formed by strands 1, 5–8, 11, and 15 (Figure 2B). This core structural motif is stabilized by interactions with β-hairpin motifs (β3–4 and β12–13), an antiparallel β-sheet running across the top of the spike (β2, β9–10, and β14), a short α-helix, and several loops that connect secondary structure elements (Figure 2B). Interestingly, electron density for residues 441–444 and 555–562 of all three chains and 609–612 of chain A was missing in our maps, indicating that these sections of the loops connecting strand 1 to strand 2, strand 9 to strand 10, and strand 13 to strand 14 may be flexible in solution (Figure 2B). The corresponding loop residues are ordered in the classical HAstV spike structures published to date, and in particular, the large loops crossing the top of the classical HAstV spike have been implicated in spike dimerization and reactivity with neutralizing antibodies.[40,50] The biological effects of these differences in HAstV MLB1 spike are currently unknown.

Despite low sequence identities, the folding topologies of the MLB and classical HAstV spikes are remarkably similar (Figures 1B and 2B). The seven-stranded β-barrel at the core of the HAstV MLB1 spike is conserved with the classical HAstV spike, with the turkey astrovirus spike,[48] and with the P2 domain of the hepatitis E virus capsid protein (Figure 2C).[40] Additionally, the MLB and classical spikes share a structural element, the β-hairpin motif (β3–4) (Figure 2B), which is also found in the turkey astrovirus spike,[48] suggesting that this structural motif may be universal in astrovirus spike domains. Most secondary structure elements, including loops, on the MLB1 spike align to an equivalent secondary structure element on the classical spike domain. However, there are notable differences in the lengths, orientations, and backbone conformations of structural elements between the two spikes. The result is a striking difference in global conformation between the MLB and classical spikes, resulting in a 3.16 Å RMSD upon alignment (Figure 2A). These observations provide a structural basis for the hypothesis that MLB HAstVs are related to—and phylogenetically distinct from—classical HAstVs.

MLB, classical, and VA astroviruses may use different receptors for cell entry

As the classical HAstV capsid spikes are implicated in viral attachment to host cells, we were interested in determining if
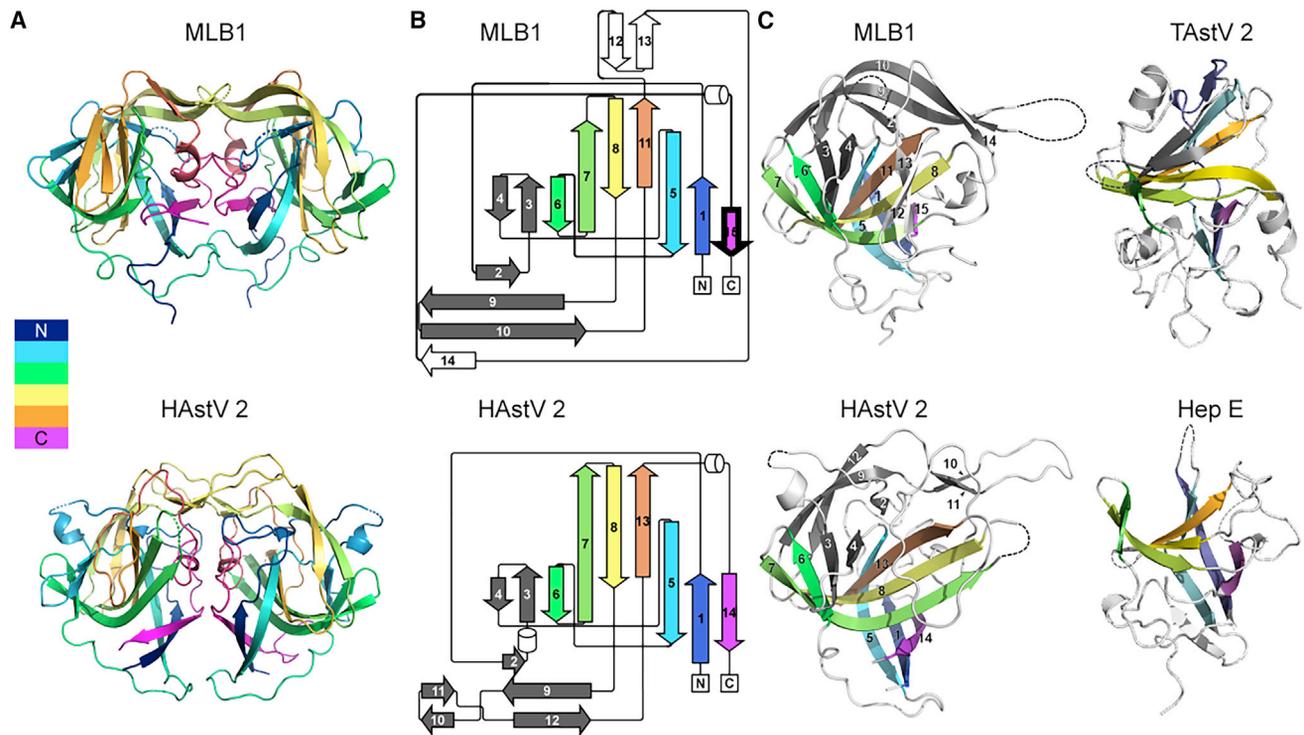
Figure 2. The classical and MLB spikes are structurally distinct but share a conserved folding topology

(A) Structures of the MLB and classical HAstV spike dimers and the results of their structural alignment. The individual spike protomers were aligned structurally, in a sequence-independent manner, using the PDBeFold server.[51] The HAstV 2 strain Oxford spike was used for the alignment and analyses in (A) and (B) (PDB: 5W1N).[49]

(B) Topology representations of the MLB and classical HAstV spike domains. The core β-barrel conserved in hepatitis E capsid P2 is colored as in Dong et al.[40] except that the most C-terminal β-strand is colored magenta instead of red. Other β-strands conserved between MLB and other known astrovirus spike domains are colored gray.

(C) Structures of the MLB and classical HAstV spike domains, the turkey astrovirus (TAstV 2) spike domain, and the hepatitis E (Hep E) P2 domain, with the core β-barrel strands colored as in (B). Flexible residues that were not included in each model are drawn approximately to scale as dashed lines.

HAstV MLBs may use the same host-cell receptor as classical and/or VA HAstVs. Receptor-binding residues have been found to be conserved in other diverse virus families.[52,53] Thus, we searched for surface-exposed patches of conserved amino acids between the classical and MLB spikes (HAstV MLB and VA spike structures cannot currently be compared as no structures of VA spikes have been published). Structural alignment of the MLB1 spike to structures of the classical HAstV serotypes 1, 2, and 8 spike domains revealed a striking lack of conserved residues, which might form a conserved receptor-binding site (Figure 3A). Importantly, we see no conservation of the putative receptor-binding residues within the P site identified previously[40] in a structural alignment of the HAstV MLB1 and HAstV 8 spikes (Figure 3B). Together with published data indicating that classical HAstV and HAstV MLB may differ both in their association with diarrhea[12,14] and their ability to propagate in Caco-2 cells,[36] our observations support a hypothesis that MLB and classical HAstV bind different receptors.

We tested this hypothesis by developing an *in vitro* HAstV infection competition assay and evaluating if recombinant MLB1 spike could inhibit infection by classical and VA HAstVs, with the rationale that if all HAstVs used the same receptor, then MLB1 spike would compete for receptors and inhibit infection. Caco-2 cells were incubated with virus on ice to allow for

attachment but not endocytosis. Cells were washed, incubated with recombinant spike protein on ice, and then incubated at 37°C to allow the viral life cycle to proceed. Infected cells were quantified using immunoperoxidase detection of virus capsid proteins. We find that classical HAstV 1 spike inhibits classical HAstV 8 infection in a dose-dependent manner, likely by competing for receptors at the cell surface (Figure 3C). In contrast, classical HAstV 1 spike had no discernable effect on HAstV VA1 infection at the concentrations tested (Figure 3C). Importantly, HAstV MLB1 spike did not inhibit infection by HAstV 8 or HAstV VA1 (Figure 3C). Unfortunately, not all potential experiments testing matching or unmatching spikes with viruses could be performed because HAstV MLB does not infect Caco-2 cells and because recombinant HAstV VA1 spike has not yet been produced. Nevertheless, these data provide evidence that two classical HAstV serotypes (HAstV 8 and HAstV 1) bind a common receptor since a spike from one classical HAstV serotype inhibits infection by another. These data also support the possibility that HAstV MLB utilizes a different receptor than both classical and VA HAstVs since MLB1 spike did not inhibit infection by either virus.

To identify putative receptor-binding residues on the MLB1 spike, we aligned MLB1-3 spike sequences and mapped conservation onto the MLB1 spike structure (Figure 3D). Several
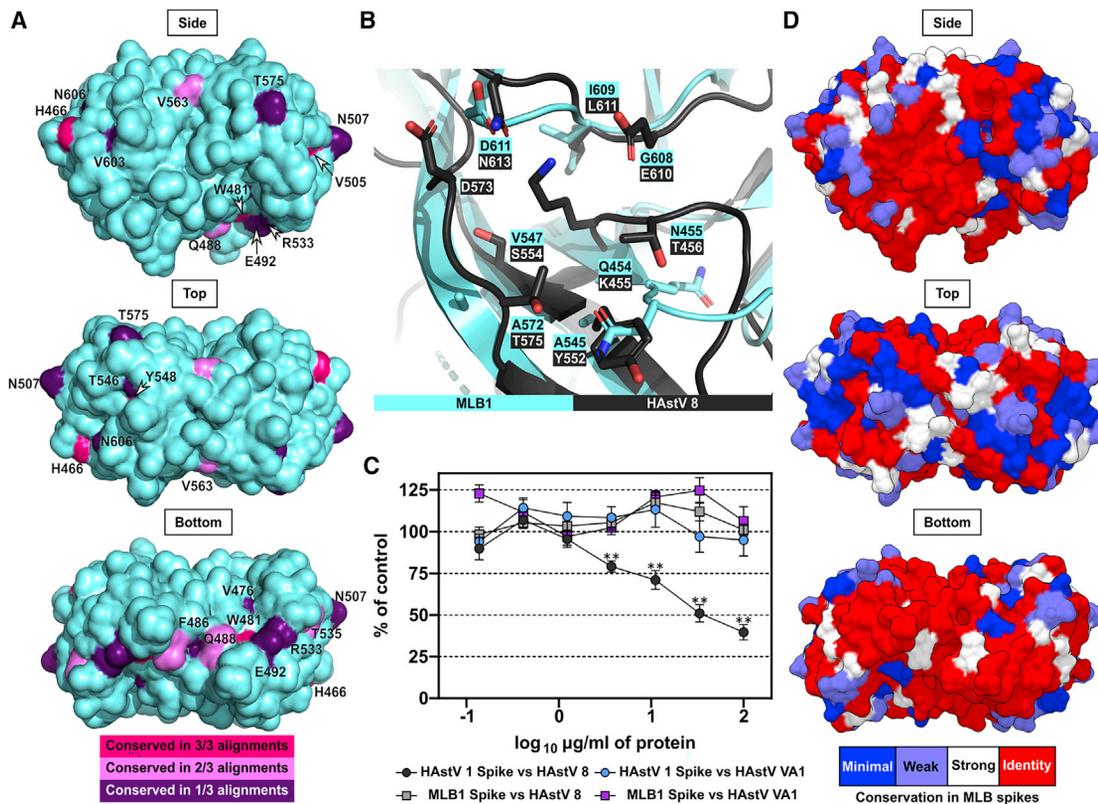
Figure 3. HAstV MLB may use a different receptor than the classical and VA1 HAstVs

(A) Sequence conservation mapped onto the MLB1 spike following structural alignment with the HAstV 8 (PDB: 3QSQ),[40] HAstV 1 (PDB: 5EWO),[38] and HAstV 2 spikes (PDB: 5W1N).[49] Identical residues that aligned structurally were colored red (for 3/3 alignments), magenta (for 2/3 alignments), or purple (for 1/3 alignments). The MLB1 spike is shown as a solvent-accessible surface. Spike structures were aligned using the PDBeFold server.[51]

(B) Comparison of the MLB1 and HAstV 8 spike amino acids in the P site. The putative receptor-binding residues in the P site[40] are shown as sticks, as are structurally aligned HAstV MLB1 residues.

(C) *In vitro* HAstV infection competition assay. Caco-2 cells were infected with HAstV in the presence of the indicated concentration of recombinant spike proteins. The MLB1 spike does not block infectivity of either HAstV 8 or HAstV VA1. The data represent the infectivity of cells infected in the presence of each recombinant spike compared with control cells infected in its absence. The arithmetic means ± SEM from three independent experiments performed in duplicate are shown. **$p < 0.01$.

(D) Conservation of MLB spike residues. All complete MLB spike sequences available in the NCBI nr database were aligned, and each position was assigned to a Clustal conservation group using UCSF Chimera.[54] Each residue in the MLB1 spike structure was then colored according to its conservation group: complete identity (red), strong conservation (white), weak conservation (light blue), and minimal conservation (dark blue).[55]

regions of high identity or similarity were observed, making it difficult to identify candidate receptor-binding site residues. The high degree of sequence conservation between MLB1-3 spikes, as well as the limited number of sequences currently available, contribute to this challenge (Figure 1B).

Retrospective structural predictions of MLB1 spike

Finally, given the recent advancements in protein structure prediction,[56,57] we sought to retrospectively ascertain whether the MLB1 spike structure could have been predicted. We first used HHpred with Modeller[58] as their server turnaround time is much faster than other servers.[59] While HHpred does not include the deep-learning methods used by other structure-prediction programs, in our specific case, HHpred with Modeller outperformed some of the methods that use deep learning, such as C-Quark[60] and both of the AlphaFold2 methods tested[56,61] (Table 2). The I-TASSER server[62]—which is from the same lab as C-Quark but does not implement deep-learning methods—was

the closest to the MLB1 spike structure by RMSD (4.19 Å) and template modeling (TM) score (0.68). RoseTTAFold[57]—the deep-learning version of Rosetta—produced the closest predictions by all-atom RMSD and local difference distance test (LDDT)[63] when the MLB spike was predicted as a monomer (0.7516 Å RMSD, 0.4211 LDDT) or a dimer (7.069 Å RMSD, 0.4278 LDDT) (Table 2). While AlphaFold2 is now able to run monomers and protein complexes, the closest it was able to get to the MLB1 spike structure using oligomers was over 11 Å RMSD—closer than the monomeric AlphaFold2 predictions (Table 2). Unsurprisingly, our efforts to solve the MLB1 spike structure by molecular replacement with these predicted models were unsuccessful.

However, during the course of peer review of this article, we submitted the MLB1 spike sequence as a target for structure prediction in the 15th Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction (CASP15). Notably, the structures predicted by the newest

Table 2. Comparison of predicted and experimentally determined MLB1 spike structures

Model comes from	RMSD[a]	All-atom RMSD[b]	TM score[c]	LDDT[d]	
I-TASSER	4.913[e]	10.399	0.67784[e]	0.357	–
RoseTTAFold Multimer	5.006	7.467	0.64681	0.4278[e]	model 1
	5.725	7.069[e]	0.6161	0.404	model 3
RoseTTAFold	5.271	7.516	0.63261	0.4211	–
HHpred/Modeller	7.098	9.898	0.65687	0.3275	–
AlphaFold2 Multimer	11.187	12.886	0.43945	0.1159	–
AlphaFold2, 48 cycles	12.496	18.725	0.42833	0.2703	–
AlphaFold2, 12 cycles	15.93	21.208	0.41335	0.236	–
C-Quark	16.388	19.328	0.29973	0.2215	–
RoseTTAFold CASP15	1.451	3.937	0.88319	0.6689	–
AlphaFold2 CASP15	1.227	7.377	0.82319	0.6011	–

[a]RMSD: root-mean-square deviation between Cα atoms, in angstroms.
[b]All-atom RMSD: root-mean-square deviation between all atoms, in angstroms.
[c]TM score: template modeling score weights smaller distance errors stronger than larger distance errors and makes the score value more sensitive to the global fold similarity than to the local structural variations.[64]
[d]LDDT: the local distance difference test is a superposition-free score that evaluates local distance differences of all atoms in a model, including validation of stereochemical plausibility.[63]
[e]The best values in each category, excluding CASP15 models.

versions of AlphaFold2 and RoseTTAFold and posted online during CASP15 were significantly improved (Table 2). In fact, we found that either of these structures could have been used successfully for molecular replacement to solve the MLB1 spike structure.

DISCUSSION

Here, we report the first structure of a divergent HAstV spike domain, the HAstV MLB1 spike. We find that the classical and MLB HAstV spikes share a common folding topology despite low sequence identities. This raises the possibility that all HAstV spikes share a conserved fold, although structures of VA-clade spikes will be needed to determine if this is the case. Outside of its folding topology, the HAstV MLB spike is strikingly different from the classical spike structures published to date. Crucially, there are no residues on the surface of the MLB spike that might clearly form a receptor-binding site that is conserved with the classical HAstVs. This suggests that the two viruses use different receptors for host-cell infection but could also be explained if MLB and classical HAstVs bind a shared receptor through different residues.

To address this ambiguity, we developed a novel *in vitro* HAstV infection competition assay and tested cross-clade inhibition of infection by recombinant HAstV spikes. We provide the first direct evidence that HAstV MLB uses a different receptor than both HAstV VA1 and the classical HAstVs. This is consistent

with published data showing that HAstV VA1 and all 8 classical HAstV serotypes—but not HAstV MLB1—grow in Caco-2 cells.[33–36] Our data raise the possibility that HAstV MLB has a very different cell tropism *in vivo* than the classical or VA1 HAstVs. This would help to explain why MLB infection has not been clearly associated with diarrhea to date.[12,14] We unfortunately could not test for inhibition of HAstV MLB infection due to the low efficiency of MLB infection in the cell lines used to date.[36] This leaves open the possibility that our MLB1 spike construct was not sufficient for receptor binding. We believe this to be unlikely, as recombinant classical HAstV spike, also produced in *E. coli*, is sufficient for cell attachment[41,42] and for inhibition of infection, as demonstrated here. Further studies are needed to confirm our findings definitively and to test their implications for cell tropism.

Having delineated the HAstV MLB spike and solved its structure also opens the door to addressing central questions in the field through experiments that were previously infeasible. The seropositivity of HAstV MLB2 and MLB3 has not been quantified to date and could not be evaluated in a serosurveillance study of HAstV MLB due to antibody cross-reactivity.[15] Using recombinant HAstV MLB1-3 spikes as antigens in seroepidemiology studies, as was performed with classical HAstV spikes,[65] should minimize cross-reactivity and may allow for the extent of HAstV MLB2 and MLB3 infection to be better understood. In a similar vein, recombinant HAstV MLB spike could be tested as an immunogen for its ability to elicit HAstV MLB-neutralizing antibodies, as has been shown for classical HAstV spikes.[50] Recombinant HAstV MLB spike could be used as a bait to identify HAstV MLB receptors in co-purification experiments.[66] Finally, reverse genetics could be used to generate chimeric HAstVs with MLB spike domains for infection studies in intestinal enteroids[25] to investigate the extent to which the spike drives cell tropism.

Given the remarkable advancements in protein structure prediction including deep-learning-based approaches,[56,57] we were surprised to find that predictions of the HAstV MLB spike structure using publicly available prediction programs were poor matches to the experimentally determined structure. One reason, particularly for methods that predicted the MLB1 spike as a monomer, could be that dimerization influences the structure of the individual protein protomers. Methods using evolutionary reasoning and contact prediction may have been especially affected by a lack of sequence information; MLB spike sequences have pairwise identities from 64% to 84% (Figure 1B), and only 27 sequences align to our MLB1 spike sequence in BLASTP searches of the NCBI nr database. Alignment of the MLB spike sequence to classical HAstV spike templates during AlphaFold2 and RoseTTAfold predictions may also have been hindered by the low sequence identities between the MLB and classical spikes. The relatively accurate models we obtained using I-TASSER and HHpred/Modeller may have come in part because these approaches use predicted secondary structure, in addition to sequence information, for template alignment.[67,68] We believe that our findings may be useful both to our colleagues developing structure prediction methods and to other structural biologists working with uncharacterized, oligomeric proteins for which limited sequence information is available. Remarkably, AlphaFold2 and RoseTTAfold predictions of the MLB1 spike generated for the CASP15 competition were very close to the

Structure
Short Article

experimental structure, so much so that the models could have been used for molecular replacement. While the latest versions of these programs are not yet publicly available, these results preview the advancements that protein structure prediction programs have made in the past year.

Overall, this study yields new insights into host-cell entry by HAstV MLB and protein structure prediction in understudied systems. Our data provide a starting point for elucidating the entry mechanisms used by nonclassical HAstVs and indicate that doing so may be important for understanding HAstV-associated diseases.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines
  - Viruses
- METHOD DETAILS
  - Sequences used for phylogenetic analysis
  - Sequences used for pairwise identity analysis
  - Production of MLB1 spike in insect cells
  - Production of MLB1 spike in *E. coli*
  - Size exclusion chromatography analysis of MLB1 spike
  - Crystallization, data collection, and structure determination
  - Purification of HAstV 1 spike
  - HAstV infection competition assay with recombinant HAstV spike proteins
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis
  - MLB spike sequences used for conservation analysis

AUTHOR CONTRIBUTIONS

Conceptualization, R.M.D. and C.F.A.; methodology, K.D.-C. and T.L.; investigation, K.D.-C. and T.L., formal analysis, F.K.; writing – original draft, K.D.-C.; writing – review & editing, R.M.D. and C.F.A.; funding acquisition, R.M.D. and C.F.A.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in their field of research or within their geographical location. One or more of the authors of this paper received support from a program designed to increase minority representation in their field of research. While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

REFERENCES

1. Finkbeiner, S.R., Li, Y., Ruone, S., Conrardy, C., Gregoricus, N., Toney, D., Virgin, H.W., Anderson, L.J., Vinjé, J., Wang, D., and Tong, S. (2009). Identification of a novel astrovirus (astrovirus VA1) associated with an outbreak of acute gastroenteritis. J. Virol. *83*, 10836–10839. https://doi.org/10.1128/JVI.00998-09.

2. Janowski, A.B. (2021). Beyond the gastrointestinal tract: the emerging and diverse tissue tropisms of astroviruses. Viruses *13*, 732. https://doi.org/10.3390/V13050732.

3. Bosch, A., Pintó, R.M., and Guix, S. (2014). Human astroviruses. Clin. Microbiol. Rev. *27*, 1048–1074. https://doi.org/10.1128/CMR.00013-14.

4. Sebire, N.J., Malone, M., Shah, N., Anderson, G., Gaspar, H.B., and Cubitt, W.D. (2004). Pathology of astrovirus associated diarrhoea in a paediatric bone marrow transplant recipient. J. Clin. Pathol. *57*, 1001–1003. https://doi.org/10.1136/JCP.2004.017178.

5. Olortegui, M.P., Rouhani, S., Yori, P.P., Salas, M.S., Trigoso, D.R., Mondal, D., Bodhidatta, L., Platts-Mills, J., Samie, A., Kabir, F., et al. (2018). Astrovirus infection and diarrhea in 8 countries. Pediatrics *141*, e20171326. https://doi.org/10.1542/PEDS.2017-1326.

6. Kapoor, A., Li, L., Victoria, J., Oderinde, B., Mason, C., Pandey, P., Zaidi, S.Z., and Delwart, E. (2009). Multiple novel astrovirus species in human stool. J. Gen. Virol. *90*, 2965–2972. https://doi.org/10.1099/VIR.0.014449-0.

7. Lanave, G., Loconsole, D., Centrone, F., Catella, C., Capozza, P., Diakoudi, G., Parisi, A., Suffredini, E., Buonavoglia, A., Camero, M., et al. (2021). Astrovirus VA1 in patients with acute gastroenteritis. Transbound. Emerg. Dis. *69*, 864–869. https://doi.org/10.1111/TBED.13979.

8. Zaraket, H., Abou-El-Hassan, H., Kreidieh, K., Soudani, N., Ali, Z., Hammadi, M., Reslan, L., Ghanem, S., Hajar, F., Inati, A., et al. (2017). Characterization of astrovirus-associated gastroenteritis in hospitalized children under five years of age. Infect. Genet. Evol. *53*, 94–99. https://doi.org/10.1016/J.MEEGID.2017.05.016.

9. Król, L., Turkiewicz, D., Nordborg, K., Englund, E., Stenberg, L., Karlsson Lindsjö, O., Lind Karlberg, M., and Pronk, C.J. (2021). Astrovirus VA1/HMO encephalitis after allogeneic hematopoietic cell transplantation: significant role of immune competence in virus control. Pediatr. Blood Cancer *68*, e29286. https://doi.org/10.1002/PBC.29286.

10. Reuter, G., Pankovics, P., and Boros, '. (2018). Nonsuppurative (aseptic) meningoencephalomyelitis associated with neurovirulent astrovirus infections in humans and animals. Clin. Microbiol. Rev. *31*, 000400-18. https://doi.org/10.1128/CMR.00040-18.

11. Vu, D.L., Bosch, A., Pintó, R.M., and Guix, S. (2017). Epidemiology of classic and novel human astrovirus: gastroenteritis and beyond. Viruses *9*, 33. https://doi.org/10.3390/V9020033.

12. Holtz, L.R., Bauer, I.K., Rajendran, P., Kang, G., and Wang, D. (2011). Astrovirus MLB1 is not associated with diarrhea in a cohort of Indian children. PLoS One *6*, e28647. https://doi.org/10.1371/JOURNAL.PONE.0028647.

13. Meyer, C.T., Bauer, I.K., Antonio, M., Adeyemi, M., Saha, D., Oundo, J.O., Ochieng, J.B., Omore, R., Stine, O.C., Wang, D., and Holtz, L.R. (2015). Prevalence of classic, MLB-clade and VA-clade astroviruses in Kenya and the Gambia. Virol. J. 12, 78. https://doi.org/10.1186/S12985-015-0299-Z.

14. Vu, D.L., Sabrià, A., Aregall, N., Michl, K., Sabrià, J., Rodriguez Garrido, V., Goterris, L., Bosch, A., Pintó, R.M., and Guix, S. (2020). A Spanish case-control study in <5 year-old children reveals the lack of association between MLB and VA astrovirus and diarrhea. Sci. Rep. 10, 1760. https://doi.org/10.1038/s41598-020-58691-3.

15. Holtz, L.R., Bauer, I.K., Jiang, H., Belshe, R., Freiden, P., Schultz-Cherry, S.L., and Wang, D. (2014). Seroepidemiology of astrovirus MLB1. Clin. Vaccine Immunol. 21, 908–911. https://doi.org/10.1128/CVI.00100-14.

16. Cordey, S., Hartley, M.A., Keitel, K., Laubscher, F., Brito, F., Junier, T., Kagoro, F., Samaka, J., Masimba, J., Said, Z., et al. (2018). Detection of novel astroviruses MLB1 and MLB2 in the sera of febrile Tanzanian children. Emerg. Microbes Infect. 7, 27. https://doi.org/10.1038/S41426-018-0025-1.

17. Cordey, S., Zanella, M.C., Wagner, N., Turin, L., and Kaiser, L. (2018). Novel human astroviruses in pediatric respiratory samples: a one-year survey in a Swiss tertiary care hospital. J. Med. Virol. 90, 1775–1778. https://doi.org/10.1002/JMV.25246.

18. Cordey, S., Vu, D.L., Schibler, M., L'Huillier, A.G., Brito, F., Docquier, M., Posfay-Barbe, K.M., Petty, T.J., Turin, L., Zdobnov, E.M., and Kaiser, L. (2016). Astrovirus MLB2, a new gastroenteric virus associated with meningitis and disseminated infection. Emerg. Infect. Dis. 22, 846–853. https://doi.org/10.3201/EID2205.151807.

19. Holtz, L.R., Wylie, K.M., Sodergren, E., Jiang, Y., Franz, C.J., Weinstock, G.M., Storch, G.A., and Wang, D. (2011). Astrovirus MLB2 viremia in febrile child - volume 17, number 11—November 2011 - emerging infectious diseases journal - CDC. Emerg. Infect. Dis. 17, 2050–2052. https://doi.org/10.3201/EID1711.110496.

20. Lau, P., Cordey, S., Brito, F., Tirefort, D., Petty, T.J., Turin, L., Guichebaron, A., Docquier, M., Zdobnov, E.M., Waldvogel-Abramowski, S., et al. (2017). Metagenomics analysis of red blood cell and fresh-frozen plasma units. Transfusion (Paris) 57, 1787–1800. https://doi.org/10.1111/TRF.14148.

21. Sato, M., Kuroda, M., Kasai, M., Matsui, H., Fukuyama, T., Katano, H., and Tanaka-Taya, K. (2016). Acute encephalopathy in an immunocompromised boy with astrovirus-MLB1 infection detected by next generation sequencing. J. Clin. Virol. 78, 66–70. https://doi.org/10.1016/J.JCV.2016.03.010.

22. Wylie, K.M., Mihindukulasuriya, K.A., Sodergren, E., Weinstock, G.M., and Storch, G.A. (2012). Sequence analysis of the human virome in febrile and afebrile children. PLoS One 7, e27735. https://doi.org/10.1371/JOURNAL.PONE.0027735.

23. Zanella, M.C., Cordey, S., Laubscher, F., Docquier, M., Vieille, G., van Delden, C., Braunersreuther, V., Ta, M.K., Lobrinus, J.A., Masouridi-Levrat, S., et al. (2021). Unmasking viral sequences by metagenomic next-generation sequencing in adult human blood samples during steroid-refractory/dependent graft-versus-host disease. Microbiome 9, 28. https://doi.org/10.1186/S40168-020-00953-3.

24. Wilson, M.R., Sample, H.A., Zorn, K.C., Arevalo, S., Yu, G., Neuhaus, J., Federman, S., Stryke, D., Briggs, B., Langelier, C., et al. (2019). Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. N. Engl. J. Med. 380, 2327–2340. https://doi.org/10.1056/NEJMOA1803396.

25. Kolawole, A.O., Mirabelli, C., Hill, D.R., Svoboda, S.A., Janowski, A.B., Passalacqua, K.D., Rodriguez, B.N., Dame, M.K., Freiden, P., Berger, R.P., et al. (2019). Astrovirus replication in human intestinal enteroids reveals multi-cellular tropism and an intricate host innate immune landscape. PLoS Pathog. 15, e1008057. https://doi.org/10.1371/JOURNAL.PPAT.1008057.

26. Gray, E.W., Angus, K.W., and Snodgrass, D.R. (1980). Ultrastructure of the small intestine in astrovirus-infected lambs. J. Gen. Virol. 49, 71–82. https://doi.org/10.1099/0022-1317-49-1-71.

27. Ingle, H., Hassan, E., Gawron, J., Mihi, B., Li, Y., Kennedy, E.A., Kalugotla, G., Makimaa, H., Lee, S., Desai, P., et al. (2021). Murine astrovirus tropism for goblet cells and enterocytes facilitates an IFN-λ response in vivo and in

28. Cortez, V., Boyd, D.F., Crawford, J.C., Sharp, B., Livingston, B., Rowe, H.M., Davis, A., Alsallaq, R., Robinson, C.G., Vogel, P., et al. (2020). Astrovirus infects actively secreting goblet cells and alters the gut mucus barrier. Nat. Commun. 11, 2097. https://doi.org/10.1038/s41467-020-15999-y.

29. Woode, G.N., Pohlenz, J.F., Gourley, N.E., and Fagerland, J.A. (1984). Astrovirus and Breda virus infections of dome cell epithelium of bovine ileum. J. Clin. Microbiol. 19, 623–630. https://doi.org/10.1128/JCM.19.5.623-630.1984.

30. Pantin-Jackwood, M.J., Spackman, E., and Day, J.M. (2008). Pathogenesis of type 2 Turkey astroviruses with variant capsid genes in 2-day-old specific pathogen free poults. Avian Pathol. 37, 193–201. https://doi.org/10.1080/03079450801932200.

31. Koci, M.D., Moser, L.A., Kelley, L.A., Larsen, D., Brown, C.C., and Schultz-Cherry, S. (2003). Astrovirus induces diarrhea in the absence of inflammation and cell death. J. Virol. 77, 11798–11808. https://doi.org/10.1128/JVI.77.21.11798-11808.2003.

32. Thouvenelle, M.L., Haynes, J.S., and Reynolds, D.L. (1995). Astrovirus infection in hatchling turkeys: histologic, morphometric, and ultrastructural findings. Avian Dis. 39, 328–336. https://doi.org/10.2307/1591875.

33. Méndez-Toss, M., Romero-Guido, P., Munguía, M.E., Méndez, E., and Arias, C.F. (2000). Molecular analysis of a serotype 8 human astrovirus genome. J. Gen. Virol. 81, 2891–2897. https://doi.org/10.1099/0022-1317-81-12-2891.

34. Janowski, A.B., Bauer, I.K., Holtz, L.R., and Wang, D. (2017). Propagation of astrovirus VA1, a neurotropic human astrovirus, in cell culture. J. Virol. 91. https://doi.org/10.1128/JVI.00740-17.

35. Brinker, J.P., Blacklow, N.R., and Herrmann, J.E. (2000). Human astrovirus isolation and propagation in multiple cell lines. Arch. Virol. 145, 1847–1856. https://doi.org/10.1007/S007050070060.

36. Vu, D.-L., Bosch, A., Pintó, R.M., Ribes, E., and Guix, S. (2019). Human astrovirus MLB replication in vitro: persistence in extraintestinal cell lines. J. Virol. 93. 005577-19. https://doi.org/10.1128/JVI.00557-19.

37. Dryden, K.A., Tihova, M., Nowotny, N., Matsui, S.M., Mendez, E., and Yeager, M. (2012). Immature and mature human astrovirus: structure, conformational changes, and similarities to hepatitis E virus. J. Mol. Biol. 422, 650–658. https://doi.org/10.1016/J.JMB.2012.06.029.

38. York, R.L., Yousefi, P.A., Bogdanoff, W., Haile, S., Tripathi, S., and DuBois, R.M. (2015). Structural, mechanistic, and antigenic characterization of the human astrovirus capsid. J. Virol. 90, 2254–2263. https://doi.org/10.1128/JVI.02666-15.

39. Toh, Y., Harper, J., Dryden, K.A., Yeager, M., Arias, C.F., Méndez, E., and Tao, Y.J. (2016). Crystal structure of the human astrovirus capsid protein. J. Virol. 90, 9008–9017. https://doi.org/10.1128/JVI.00694-16.

40. Dong, J., Dong, L., Méndez, E., and Tao, Y. (2011). Crystal structure of the human astrovirus capsid spike. Proc. Natl. Acad. Sci. USA 108, 12681–12686. https://doi.org/10.1073/PNAS.1104834108.

41. Ricemeyer, L., Aguilar-Hernández, N., López, T., Espinosa, R., Lanning, S., Mukherjee, S., Cuellar, C., López, S., Arias, C.F., and DuBois, R.M. (2022). Structures of two human astrovirus capsid/neutralizing antibody complexes reveal distinct epitopes and inhibition of virus attachment to cells. J. Virol. 96, e0141521. https://doi.org/10.1128/jvi.01415-21.

42. Bogdanoff, W.A., Campos, J., Perez, E.I., Yin, L., Alexander, D.L., and DuBois, R.M. (2017). Structure of a human astrovirus capsid-antibody complex and mechanistic insights into virus neutralization. J. Virol. 91, 018599-16. https://doi.org/10.1128/jvi.01859-16.

43. Méndez, E., Muñoz-Yañez, C., S'nchez-San Martín, C., Aguirre-Crespo, G., Baños-Lara, M.d.R., Gutierrez, M., Espinosa, R., Acevedo, Y., Arias, C.F., and López, S. (2014). Characterization of human astrovirus cell entry. J. Virol. 88, 2452–2460. https://doi.org/10.1128/JVI.02908-13.

44. Méndez, E., Salas-Ocampo, E., and Arias, C.F. (2004). Caspases mediate processing of the capsid precursor and cell release of human astroviruses. J. Virol. 78, 8601–8608. https://doi.org/10.1128/JVI.78.16.8601-8608.2004.

45. Banos-Lara, M.d.R., and Méndez, E. (2010). Role of individual caspases induced by astrovirus on the processing of its structural protein and its release from the cell through a non-lytic mechanism. Virology *401*, 322–332. https://doi.org/10.1016/J.VIROL.2010.02.028.

46. Lee, T.W., and Kurtz, J.B. (1981). Serial propagation of astrovirus in tissue culture with the aid of trypsin. J. Gen. Virol. *57*, 421–424. https://doi.org/10.1099/0022-1317-57-2-421.

47. Aguilera-Flores, C., López, T., Zamudio, F., Sandoval-Jaime, C., Pérezc, E.I., López, S., DuBois, R., and Arias, C.F. (2022). The capsid precursor protein of astrovirus VA1 is proteolytically processed intracellularly. Preprint at bioRxiv. https://doi.org/10.1101/2022.04.28.489973.

48. DuBois, R.M., Freiden, P., Marvin, S., Reddivari, M., Heath, R.J., White, S.W., and Schultz-Cherry, S. (2013). Crystal structure of the avian astrovirus capsid spike. J. Virol. *87*, 7853–7863. https://doi.org/10.1128/JVI.03139-12.

49. Bogdanoff, W.A., Perez, E.I., López, T., Arias, C.F., and DuBois, R.M. (2018). Structural basis for escape of human astrovirus from antibody neutralization: broad implications for rational vaccine design. J. Virol. *92*, 015466-17. https://doi.org/10.1128/JVI.01546-17.

50. Espinosa, R., López, T., Bogdanoff, W.A., Espinoza, M.A., López, S., DuBois, R.M., and Arias, C.F. (2019). Isolation of neutralizing monoclonal antibodies to human astrovirus and characterization of virus variants that escape neutralization. J. Virol. *93*, 014655-18. https://doi.org/10.1128/JVI.01465-18.

51. Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. *60*, 2256–2268. https://doi.org/10.1107/S0907444904026460.

52. Shi, Y., Wu, Y., Zhang, W., Qi, J., and Gao, G.F. (2014). Enabling the "host jump": structural determinants of receptor-binding specificity in influenza A viruses. Nat. Rev. Microbiol. *12*, 822–831. https://doi.org/10.1038/nrmicro3362.

53. Choi, J.M., Hutson, A.M., Estes, M.K., and Prasad, B.V.V. (2008). Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. Proc. Natl. Acad. Sci. USA *105*, 9175–9180. https://doi.org/10.1073/PNAS.0803275105.

54. Meng, E.C., Pettersen, E.F., Couch, G.S., Huang, C.C., and Ferrin, T.E. (2006). Tools for integrated sequence-structure analysis with UCSF Chimera. BMC Bioinf. *7*, 339. https://doi.org/10.1186/1471-2105-7-339.

55. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. *31*, 3497–3500. https://doi.org/10.1093/NAR/GKG500.

56. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

57. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., Wang, J., Cong, Q., Kinch, L.N., Schaeffer, R.D., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. Science *373*, 871–876. https://doi.org/10.1126/SCIENCE.ABJ8754.

58. Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A.N., and Alva, V. (2018). A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J. Mol. Biol. *430*, 2237–2243. https://doi.org/10.1016/J.JMB.2017.12.007.

59. Haas, J., Barbato, A., Behringer, D., Studer, G., Roth, S., Bertoni, M., Mostaguir, K., Gumienny, R., and Schwede, T. (2018). Continuous automated model evaluation (CAMEO) complementing the critical assessment of structure prediction in CASP12. Proteins *86*, 387–398. https://doi.org/10.1002/PROT.25431.

60. Mortuza, S.M., Zheng, W., Zhang, C., Li, Y., Pearce, R., and Zhang, Y. (2021). Improving fragment-based ab initio protein structure assembly using low-accuracy contact-map predictions. Nat. Commun. *12*, 5011. https://doi.org/10.1038/s41467-021-25316-w.

61. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., et al. (2022). Protein complex prediction with AlphaFold-Multimer. Preprint at bioRxiv. https://doi.org/10.1101/2021.10.04.463034.

62. Zheng, W., Zhang, C., Bell, E.W., and Zhang, Y. (2019). I-TASSER gateway: a protein structure and function prediction server powered by XSEDE. Future Gener. Comput. Syst. *99*, 73–85. https://doi.org/10.1016/J.FUTURE.2019.04.011.

63. Mariani, V., Biasini, M., Barbato, A., and Schwede, T. (2013). lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. Bioinformatics *29*, 2722–2728. https://doi.org/10.1093/BIOINFORMATICS/BTT473.

64. Zhang, Y., and Skolnick, J. (2004). Scoring function for automated assessment of protein structure template quality. Proteins *57*, 702–710. https://doi.org/10.1002/PROT.20264.

65. Meyer, L., Delgado-Cunningham, K., Lorig-Roach, N., Ford, J., and Dubois, R.M. (2021). Human astrovirus 1–8 seroprevalence evaluation in a United States adult population. Viruses *13*, 979. https://doi.org/10.3390/V13060979.

66. Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature *426*, 450–454. https://doi.org/10.1038/nature02145.

67. Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. *5*, 725–738. https://doi.org/10.1038/NPROT.2010.5.

68. Gabler, F., Nam, S.Z., Till, S., Mirdita, M., Steinegger, M., Söding, J., Lupas, A.N., and Alva, V. (2020). Protein sequence analysis using the MPI bioinformatics toolkit. Curr. Protoc. Bioinf. *72*, e108. https://doi.org/10.1002/CPBI.108.

69. Janowski, A.B., Owen, M.C., Dudley, H., López, T., Espinosa, R., Elvin-Lewis, M., Colichon, A., Arias, C.F., Burbelo, P.D., and Wang, D. (2021). High seropositivity rate of neutralizing antibodies to astrovirus VA1 in human populations. mSphere *6*, e0048421. https://doi.org/10.1128/MSPHERE.00484-21.

70. Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132. https://doi.org/10.1107/S0907444909047337.

71. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol. *75*, 861–877. https://doi.org/10.1107/S2059798319011471.

72. Terwilliger, T.C., Adams, P.D., Read, R.J., McCoy, A.J., Moriarty, N.W., Grosse-Kunstleve, R.W., Afonine, P.v., Zwart, P.H., and Hung, L.W. (2009). Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. Acta Crystallogr. D Biol. Crystallogr. *65*, 582–601. https://doi.org/10.1107/S0907444909012098.

73. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674. https://doi.org/10.1107/S0021889807021206.

74. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501. https://doi.org/10.1107/S0907444910007493.

75. Aguilar-Hern'ndez, N., Meyer, L., López, S., Dubois, R.M., and Arias, C.F. (2020). Protein disulfide isomerase A4 is involved in genome uncoating during human astrovirus cell entry. Viruses *13*, 53. https://doi.org/10.3390/V13010053.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal serum to HAstV 8 (Yuc8)	Méndez et al., 2004	N/A
Monoclonal antibody MAb 2A2	Janowski et al., 2021	N/A
Bacterial and virus strains		
Bacterial cells: *E. coli* strain T7 Express	New England Biolabs	Cat#C2566H
Bacterial cells: *E. coli* strain B843 (DE3) pLysS	Novagen	Cat#69042
HAstV 8 strain Yuc8	Méndez et al., 2004	N/A
HAstV VA1	Janowski et al., 2017	N/A
Critical commercial assays		
flashBAC Baculovirus Expression System	Mirus Bio	Cat#MIR6115
Deposited data		
MLB1 Spike structure	This paper	PDB: 7UZT
Experimental models: Cell lines		
Insect cells: *Spodoptera frugiperda* Sf9 cell line IPLB-Sf-21-AE	Expression Systems	Cat#94-001S
Caco-2 cells: C2Bbe1 cells, derived from the colon adenocarcinoma Caco-2 cell line	American Type Culture Collection	Cat# CRL-2102
Oligonucleotides		
pBacPAK8-MLB1-spike_Forward1: ccaccatcaTCATTCAGAGACAACCTACACTGGCC	This paper	N/A
pBacPAK8-MLB1-spike_Forward2: gcagatgcaccaccaccatcatcaccaccatcaTCATTCAGAGACAACC	This paper	N/A
pBacPAK8-MLB1-spike_Reverse: gcaccagagcgagctctTTAGCGGTGCGGTATTCCCTC	This paper	N/A
pRSFDuet-MLB1-spike_Forward: ctggttccgcgtggtTCATCAGAGACAACCTACACTGGCC	This paper	N/A
pRSFDuet-MLB1-spike_Reverse: GCACCAGAGCGAGCTCT	This paper	N/A
Recombinant DNA		
Codon-optimized synthetic gene encoding residues 399-653 of the MLB1 capsid protein (NCBI YP_002290968.1)	Genewiz	N/A
pBacPAK8	Takara Bio	631402
pRSFDuet	Novagen	71341
Software and algorithms		
PyMol Molecular Graphics	System Schrodinger, LLC	https://pymol.org/2/
TM-align	Zhang Lab, University of Michigan	https://zhanggroup.org/TM-align/
lDDT (Local Distance Difference Test)	University of Basel	https://swissmodel.expasy.org/lddt

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Rebecca DuBois (rmdubois@ucsc.edu).

Materials availability
Plasmids produced in this study are available from the lead contact.

Structure
Short Article

Data and code availability

Coordinates and structure factors for the MLB1 spike structure have been deposited in the Protein Data Bank and are publicly available as of the date of publication with the accession code 7UZT. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

Recombinant MLB1 spike protein was purified from the expression in insect cells, *Spodoptera frugiperda* Sf9 cell line IPLB-Sf-21-AE (Expression Systems), or in bacterial cells, *E. coli* strain T7 Express (New England Biolabs) or *E. coli* strain B843 (DE3) pLysS (Novagen). C2Bbe1 cells, derived from the colon adenocarcinoma Caco-2 cell line, were obtained from the American Type Culture Collection, and propagated in a 10% CO_2 atmosphere at 37°C in Dulbecco's modified Eagle's medium (DMEM)-reduced serum (Thermo Scientific HyClone, Logan, UT) supplemented with 5% fetal bovine serum (FBS) (Cansera, Ontario, Canada).

Viruses

HAstV 8 strain Yuc8 was isolated in our laboratory, and it was grown as described previously.[44] HAstV VA1 was obtained from David Wang's laboratory[34]; a virus stock was produced in C2Bbe1 cells infecting at a MOI of 0.1, and collecting the virus at 4 days post-infection by lysing the cells with three of freeze-thaw cycles.[44,69]

METHOD DETAILS

Sequences used for phylogenetic analysis

The cladogram in Figure 1A was constructed using complete capsid protein sequences with the following accession numbers: Human astrovirus 1: AAC34717.1 (NCBI), Human astrovirus 2: AAA62427.1 (NCBI), Human astrovirus 3: Q9WFZ0 (uniprot), Human astrovirus 4: Q3ZN05 (uniprot), Human astrovirus 5: Q4TWH7 (uniprot), Human astrovirus 6: AZB52207.1 (NCBI), Human astrovirus 7: Q96818 (uniprot), Human astrovirus 8: AAF85964.1 (uniprot), Human astrovirus MLB1: YP_002290968.1 (NCBI), Human astrovirus MLB2: YP_004934010.1 (NCBI), Human astrovirus MLB3: YP_006905854.1 (NCBI), Human astrovirus VA1: ADH93577.1 (NCBI), Human astrovirus VA2: ACX83591.2 (NCBI), Human astrovirus VA3: YP_006905860.1 (NCBI), Human astrovirus VA4: YP_006905857.1 (NCBI), Human astrovirus VA5: AJI44022.1 (NCBI), Turkey astrovirus 1: NP_059949.1 (NCBI), Turkey astrovirus 2: NP_987088.1 (NCBI), Turkey astrovirus 3: AAV37187.1 (NCBI).

Sequences used for pairwise identity analysis

Pairwise identities of the following astrovirus spike and core domain sequences were used in Figure 1B: Human astrovirus 1: residues 431-644 (spike) and 80-412 (core), Human astrovirus 2: residues 429-644 (spike) and 77-411 (core), Human astrovirus 3: residues 432-645 (spike) and 79-414 (core), Human astrovirus 4: residues 430-644 (spike) and 77-412 (core), Human astrovirus 5: residues 429-641 (spike) and 76-411 (core), Human astrovirus 6: residues 430-642 (spike) and 77-412 (core), Human astrovirus 7: 431-644 (spike) and 78-413 (core), Human astrovirus 8: residues 490-705 (spike) and 137-472 (core), Human astrovirus MLB1: residues 420-646 (spike) and 62-397 (core), Human astrovirus MLB2: residues 417-643 (spike) and 62-395 (core), Human astrovirus MLB3: residues 417-643 (spike) and 62-395 (core), Human astrovirus VA1: residues 394-697 (spike) and 69-390 (core), Human astrovirus VA2: residues 404-688 (spike) and 69-392 (core), Human astrovirus VA3: residues 388-691 (spike) and 63-384 (core), Human astrovirus VA4: residues 408-685 (spike) and 69-392 (core), Human astrovirus VA5: residues 406-678 (spike) and 68-391 (core). Accession numbers match those used for phylogenetic analysis.

Production of MLB1 spike in insect cells

A codon-optimized synthetic gene encoding residues 399-653 of the MLB1 capsid protein (NCBI YP_002290968.1) was cloned into the plasmid pBacPAK8 in frame with an N-terminal 10-histidine tag. A recombinant baculovirus stock was generated using the flashBAC system (Mirus bio). Sf9 insect cells (Expression Systems) in ESF-921 media were infected at a density of 2 million viable cells/mL with 0.025 mL of baculovirus stock/mL and cultured at 180 rpm, 27°C. Cells were harvested by centrifugation at 4 days post-infection, re-suspended in buffer A (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 20 mM imidazole, 2 mM $MgCl_2$, protease inhibitor cocktail (Millipore), and benzonase), and lysed via microfluidization. The lysate was clarified by centrifugation (40,000 x g), 0.22-um-filtered, and MLB1 spike was purified from the supernatant using TALON metal affinity chromatography. Fractions containing the MLB1 spike were dialyzed into 20 mM Tris-HCl pH 8.0, 20 mM NaCl and further purified using anion exchange chromatography. Fractions containing the MLB1 spike were dialyzed into 10 mM Tris-HCl pH 8.0, 150 mM NaCl, concentrated to ∼ 5 mg/mL, flash frozen in liquid nitrogen, and stored indefinitely at -80°C.

Production of MLB1 spike in *E. coli*

The MLB1 spike gene was cloned into the pRSFDuet plasmid in frame with an N-terminal 10-histidine tag, an alanine-serine linker, and a thrombin cleavage site. The plasmid was transformed into *E. coli* strain T7 Express (New England Biolabs) and grown in LB/kanamycin, and expression was induced with 1 mM IPTG at 18°C for 16 hours. For selenomethionine-substituted protein, the plasmid was transformed into *E. coli* strain B843 (DE3) pLysS (Novagen) and grown in M9 minimal medium supplemented with kanamycin, chloramphenicol,

and 0.04 mg/mL selenomethionine. Cells were harvested by centrifugation, re-suspended in buffer A, and lysed using ultrasonication. The lysate was clarified by centrifugation (40,000 x g), 0.22-um-filtered, and MLB1 spike was purified from the supernatant using TALON metal affinity chromatography. For crystallography, fractions containing selenomethionine-substituted MLB1 spike were dialyzed into 10 mM Tris-HCl pH 8.0, 150 mM NaCl, and concentrated and frozen as described above. For HAstV-inhibition experiments, fractions containing the MLB1 spike were dialyzed into phosphate-buffered saline pH 7.65, concentrated to $\sim$ 1 mg/mL, and frozen as described above.

Size exclusion chromatography analysis of MLB1 spike
Purified MLB1 spike was analyzed by size exclusion chromatography with a Superdex 200 increase 10/300 column. The oligomeric state of MLB1 spike was estimated by comparing its retention volume to those of the proteins in Gel Filtration Standards (Bio-Rad).

Crystallization, data collection, and structure determination
Native MLB1 spike at 5 mg/mL was crystallized in 100 mM Bis-Tris HCl pH 5.5, 350 mM $MgCl_2$, 22.5% w/v PEG 3350 using the hanging drop method. Selenomethionine-substituted MLB1 spike at 5 mg/mL was crystallized in 100 mM Bis-Tris HCl pH 5.5, 300 mM $MgCl_2$, 21% w/v PEG 3350 using the hanging drop method. Native spike crystals were cryoprotected using 100 mM Bis-Tris HCl pH 5.5, 300 mM $MgCl_2$, 25% w/v PEG 3350, and 18% v/v EDG (1:1:1 ethylene glycol, DMSO, and glycerol). Selenomethio-nine-substituted spike crystals were cryoprotected using a solution which instead had 200 mM $MgCl_2$ but was otherwise identical. All crystals were flash-frozen in liquid nitrogen. Data were collected at cryogenic temperatures at ALS beamline 5.0.2 and were processed using XDS.[70] The MLB1 spike structure was solved using SAD and an initial model was autobuilt using PHENIX AutoSol.[71,72] This model was then used to phase a higher-resolution native dataset using molecular replacement with PHASER.[73] The final model was refined and manually built using PHENIX and COOT,[74] respectively. Coordinates and structure factors for the MLB1 spike structure have been deposited in the Protein Data Bank with the accession code 7UZT.

Purification of HAstV 1 spike
The spike domain of classical human astrovirus serotype 1 was produced in *E. coli* and purified as described previously.[38] Briefly, cDNA encoding HAstV-1 capsid protein residues 429 to 645 (Accession# AAC34717.1) was cloned into pET52b in frame with a C-terminal thrombin cleavage site and 10-histidine purification tag. The plasmid was transformed into *E. Coli* strain BL21 (DE3) and expression was induced with 1mM IPTG at 18ºC for 16 h. Cell pellets were resuspended in Buffer A and lysed by ultrasonication. The lysate was clarified by centrifugation (40,000 x g), 0.22-um-filtered, and HAstV-1 spike was purified HisTrap metal-affinity chromatography. Fractions containing HAstV-1 spike were dialyzed into phosphate-buffered saline pH 7.65, concentrated to $\sim$ 1 mg/mL, and flash frozen as described above.

HAstV infection competition assay with recombinant HAstV spike proteins
C2Bbe1 cells were grown in 96-well plates until confluence, the growth medium was removed and replaced by MEM pre-cooled to 4°C, and the cells were incubated for 20 min on ice. After this time the MEM was replaced by virus diluted in MEM at an MOI of 0.02 and incubated on ice for 1 h; the unbound virus was washed away with MEM and the indicated concentration of the purified HAstV spike proteins, diluted in MEM, was added and incubated for 1 h on ice. The plates were then transferred to 37°C for 1 h, washed with MEM to remove the proteins, and then were incubated in DMEM supplemented with non-essential amino acids (NEAA) for 18 h for HAstV 8 (Yuc8), or for 48 h for HAstV VA1. After this time, the cells were processed by an immunoperoxidase assay to detect the infected cells, as previously described,[75] with some modifications. Briefly, the cells were fixed at room temperature for 20 minutes with 2% formaldehyde diluted in PBS and permeabilized with 0.2% Triton X-100-PBS for 15 min. To stain the cells infected with HAstV 8, a rabbit polyclonal serum to HAstV 8 (Yuc8) was used,[44] and for cells infected with HAstV VA1 we employed the monoclonal antibody MAb 2A2.[69] Three independent experiments were performed in duplicate.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis
Statistical significance of the HAstV infection competition assay data was evaluated by comparing the infectivity in each protein dilution, compared to control cells without protein, using the Mann–Whitney test with GraphPad Prism 5 (GraphPad Software Inc., San Diego, CA, USA).

MLB spike sequences used for conservation analysis
The following sequences were used for the analysis of MLB-spike conservation in Figure 3D: NCBI accession numbers ASU91602.1 residues 345-571, QCC21357.1 residues 420-646, QAA77562.1 residues 467-693, BAN57337.1 residues 420-646, ADJ38388.1 residues 420-646, AFD61563.1 residues 420-646, BAU68081.1 residues 420-646, BAN57334.1 residues 420-646, ACN44171.1 res-idues 420-646, UAW96012.1 residues 420-646, AGZ15312.1 residues 420-646, YP_006905854.1 residues 417-643, AJC68713.1 residues 417-643, BAN62843.1 residues 417-643, AZU90755.1 residues 417-643, YP_004934010.1 residues 417-643, QCC21360.1 residues 417-643, APB03099.1 residues 417-643, AER41415.1 residues 417-643, AMR45107.1 residues 417-643, and YP_002290968.1 residues 399-653.