Chapter 3
Astrovirus Structure and Assembly

Rebecca M. Dubois, Kelly A. Dryden, Mark Yeager, and Yizhi J. Tao

Abstract Recent structural studies on the astrovirus virion and viral proteins have yielded exciting new insights into the molecular mechanisms of the astrovirus life cycle. The 25 Å-resolution cryo-electron microscopy (Cryo-EM) reconstructions of the astrovirus virion reveal a solid capsid shell studded with spikes. Proteolytic maturation of the virus particle results in capsid conformational changes, most prominently at the spikes. High-resolution crystal structures of the human and avian astrovirus capsid spike domains have shed light on potential host receptors and species specificity. Together, both the structural studies on the astrovirus virion and capsid spike domains have revealed similarities to hepatitis E virus, suggesting an evolutionary relationship. The only other structural information on astrovirus is from the high-resolution crystal structure of the protease that is involved in nonstructural polyprotein processing. Overall, these structural studies have led a better understanding of the astrovirus life cycle, including astrovirus assembly, virus release, maturation, receptor binding, antibody neutralization, and nonstructural polyprotein processing.

R.M. Dubois, Ph.D. (✉)
Department of Structural Biology, St. Jude Children's Research Hospital,
262 Danny Thomas Place, Memphis, TN 38105, USA
e-mail: rebecca.dubois@stjude.org

K.A. Dryden, Ph.D.
Department of Molecular Physiology and Biological Physics, University of Virginia,
Snyder Building, Room 232, Charlottesville, VA 22903, USA
e-mail: kdryden@virginia.edu

M. Yeager, M.D., Ph.D.
Department of Molecular Physiology and Biological Physics, University of Virginia,
1340 Jefferson Park Avenue, Charlottesville, VA 22908, USA
e-mail: yeager@virginia.edu

Y.J. Tao, Ph.D.
Department of Biochemistry and Cell Biology, Rice University,
6100 Main Street, MS140, Houston, TX 77005, USA
e-mail: ytao@rice.edu

S. Schultz-Cherry (ed.), *Astrovirus Research: Essential Ideas, Everyday Impacts, Future Directions*, DOI 10.1007/978-1-4614-4735-1_3,

Introduction

Structural studies on the astrovirus virion and viral proteins are important to understand the molecular details of the astrovirus life cycle. Recent structural studies of the astrovirus virion using cryo-electron microscopy have yielded important insights into the proteolytic maturation and corresponding conformational changes in the astrovirus virion. Additionally, recent high-resolution X-ray crystallographic studies on human and avian astrovirus capsid spikes have revealed insights into receptor binding, host cell entry, and species specificity. Together, the studies on the astrovirus virion and capsid spike have uncovered striking structural similarities to hepatitis E virus, suggesting that these viruses have a closer phylogenetic relationship than previously recognized. Finally, the high-resolution X-ray crystallographic structure of the astrovirus protease revealed a number of distinctive features that provide insight into its role in nonstructural polyprotein processing events. Overall, astrovirus structural studies provide a foundation for future mutagenesis studies to probe protein functions and roles in the astrovirus life cycle and pathogenesis. Furthermore, these structural studies will have a large impact on the development of antivirals and vaccines targeting astrovirus.

Cryo-EM Structure of Immature and Mature Human Astrovirus

Astroviruses are assembled from the structural protein VP90, an 87–90 kDa capsid precursor protein, which is encoded by the 3′-terminal one-third of the genome (ORF-2). This capsid protein has a highly conserved N-terminal domain and hypervariable C-terminal domain and must undergo a cascade of cleavages to become infectious. The final size and number of structural proteins that form the capsid vary depending on the serotype (reviewed by Krishna [23]). VP90 undergoes an initial caspase cleavage at the C-terminus to generate VP70 (a 70–79 kDa protein), which is required for particle release from cells [28]. Viral infectivity is further dependent upon trypsin cleavage and the genesis, usually, of three structural proteins, VP34, VP27 (27–29 kDa), and VP25 (25–26 kDa) [4, 27, 33]. The current consensus is that the cleavage product VP34 contains the conserved, N-terminal domain, while VP27 and VP25 contain the variable domain, which mediate the attachment of virions to cell surface receptors and contain epitopes for neutralizing antibodies [31]. VP25 is generated upon additional trimming at the N-terminus of VP27.

Astroviruses were initially named for the distinctive star-like appearance of the viral surface that was seen in ~10% of the fecally shed viral particles evaluated by negative-stain electron microscopy (EM) [2, 11]. EM ultrastructural analysis of infectious HAstV generated in cell culture showed spherical particles with a surface studded with spikes and an external diameter of 41 nm [30]. The characteristic star-like surface appearance was inducible after brief alkaline (pH 10) treatment of the particles, which made them less stable, and suggested that particles with star-like morphology may represent an intermediate step toward viral disassembly. More recently, the structures of immature (VP70) and mature virions have been calculated

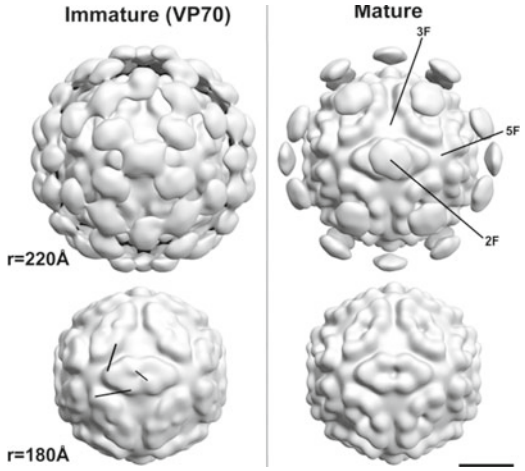

Fig. 3.1 Cryo-EM reconstructions of immature and mature HAstV. Immature (*left*) and mature (*right*) particles have very similar capsids as displayed by truncating the maps at radius 180 Å (*lower*), while the mature particles lack 60 of the 90 spikes visible on immature virions (*upper*). Symmetry axes are indicated on mature virion. One set of probable subunit contributors to the dimer spikes is displayed on the immature capsid. *Bar* = 100 Å

to moderate 25 Å-resolution by electron cryomicroscopy (cryo-EM) and image analysis [14]. The majority of the variable domain (a.a. 425–646) has been crystallized [13] and is discussed in detail later in this chapter.

Structural Characteristics and Similarity to Hepatitis E Virus

The cryo-EM reconstructions of immature assemblies of VP70 and mature, infectious particles both display two layers of density: an inner, continuous capsid layer and an outer layer composed of disconnected or weakly connected globular densities (Fig. 3.1). The capsids of the immature and mature particles are nearly identical at this resolution. The shells are ~350 Å in diameter and ~45 Å thick with features that support organization as $T = 3$ icosahedral symmetry, therefore, assembled from 180 protein subunits. They have a solid, chiseled appearance forming triangular plateaus with the edges along the icosahedral twofold axes in close contact (Fig. 3.1, lower). There are depressions at the fivefold and threefold symmetry axes and a groove across the local twofold axes. The highly conserved N-terminal domain has a predicted secondary structure that would be homologous to the canonical β-sandwich fold in many other icosahedral viruses [23]. Therefore, the HAstV capsid shell is attributed to the conserved domain, which primarily comprises VP34 in the mature virions.

The distal layer of density is composed of discrete domains and represents the spikes that have been observed directly by electron microscopy (Fig. 3.1, upper). These globular domains have very poorly defined connections to the underlying capsid, implying that they may be flexible or have thin linkers. The densities are ~40 Å in diameter, although more elliptical than spherical, and extend to an outer radius of ~215 Å making the overall diameter of the particles ~430 Å. These densities are attributed the C-terminal, variable domain. Details of the astrovirus spike structure have been further elucidated from the crystal structure of the dimerized VP25 fragment [13].

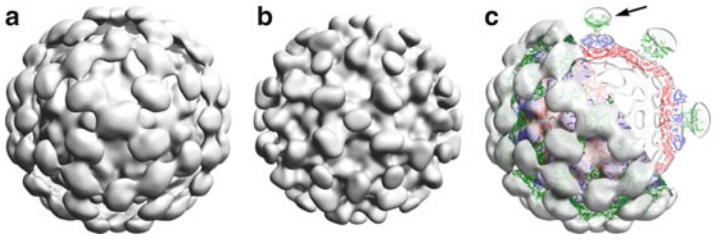

Fig. 3.2 Comparison of HAstV and HEV structures. The tertiary structure of immature HAstV (**a**) is very similar to that of HEV (**b**). The HEV-lp map (emd ID: 5,173) was truncated to 25 Å resolution for consistency. (**c**) The $T=3$ atomic model (PDB ID: 3IYO) docked into the HAstV map with no scaling. Chain is color coded for the HEV S- (*red*), M- (*blue*), and P-domains (*green*). The greatest difference between the maps is the greater radius of the HAstV spikes (*arrow*)

The size, shape, and architecture of immature HAstV is also remarkably similar to that of a $T=3$ hepatitis E virus-like particle (HEV-lp) calculated by cryo-EM [36] (Fig. 3.2a, b). Like astrovirus, HEV is a small, nonenveloped, single-stranded, positive-sense RNA virus with a genome size of about 7.2 kb. At present, it is the sole member of the *Hepeviridae* family. In addition to the cryo-EM reconstruction, crystal structures have also been determined for the HEV $T=1$ subviral particles assembled from recombinant proteins [19, 37]. The HEV structural protein has been described by sequential domains: a positively charged N-terminus likely involved in RNA packaging, a conserved β-sandwich forming the shell (S-domain), two protruding densities (P1- and P2-domains), and a C-terminal region that must be removed for in vitro assembly (the P1- and P2-domains are also referred to as the M- and P-domains, respectively). The $T=1$ crystal structure of HEV allowed detailed modeling of the protein subunits into the $T=3$ cryo-EM map. It was proposed that the capsid is assembled from dimers around the fivefold axes and across the icosahedral twofold axes. HEV using dimers as building blocks is further supported by the requirement for dimerization of the P2-domain for both assembly and host cell interactions [25, 37].

Like HEV, the overall structural organization of HAstV polyprotein can be similarly described, with the astrovirus conserved domain corresponding to the arginine-rich and S-domain of HEV, and the variable domain corresponding to the P1-, P2-, and cleaved C-terminal domains of HEV. Sequence alignment of the two viruses has 21, 13, and 14% identity for the S-, P1-, and P2-domains, respectively [14]. The major noticeable difference between the HAstV and HEV structures is that the astrovirus spikes extend to a greater radius than those of HEV, which has more compact spikes and an outer radius of ~190 Å (Fig. 3.2c, arrow). The HEV atomic model docks exceedingly well into the HAstV cryo-EM density maps without any scaling (Fig. 3.2c). The subunit organization of astrovirus can be inferred by analogy, and Dong et al. used the HEV atomic model to build a corresponding model for HAstV [13]. As with HEV, the spike stoichiometry and orientation suggests that they are dimers with contributions from subunits situated on different trimeric facets (Fig. 3.1, lower left). Using the nomenclature of HEV, AB dimers

encircle the fivefold axes, while neighboring C subunits form dimers across the icosahedral twofold axes. In the case of HEV, and perhaps more so for HAstV, this arrangement leads to different interactions for the peripentonal dimers versus the dimers on the twofold axes. The spike bases for the twofold dimers are in much closer proximity than those around the fivefold, which are situated at nearly twice the distance across the local twofold groove. Extrapolating further from the HEV structure, it is likely that there is an underlying difference in the hinge angles between these dimers, which may reflect the presence or absence of RNA in close contact on the interior surface [36].

Partial Loss of Visible Spikes upon Trypsin Cleavage

While the capsid shells of both the immature and mature HAstV particles are essentially the same, there is a dramatic loss of spikes during trypsin cleavage (Fig. 3.1, upper). The reconstruction of the immature, uncleaved particles has 90 globular, or spike, densities. In the reconstruction of the mature particles, however, only the 30 spike densities located at the icosahedral twofold axes remain visible. Trypsin cleavage of the immature particles produces three products: we presume that VP34 forms the capsid shell and the two different C-terminal products, VP27 and VP25, form the surface spikes. Although a first assumption would be that the trypsin cleavage has removed two-thirds of the spikes from the particles, previously published SDS-PAGE gels show that the amount of VP34 present in infectious particles was comparable to that of VP27 and VP25 [4, 27, 33], indicating that most of the VP27 and VP25, although covalently independent, remain associated with the particle. Therefore, an alternative explanation is that the stoichiometry of the spikes may be conserved, and trypsin cleavage results in conformational flexibility so that the density is lost upon image averaging.

Proteolytic cleavage is a common mechanism for activation in both enveloped and nonenveloped viruses, including influenza virus [21, 24], coronavirus [15], rotavirus [16], reovirus [7], and alphaviruses [32]. A recurrent theme is that cleavage triggers a conformational change in a surface spike or other cell-attachment domain. In the case of rotavirus, the spike is initially completely unstructured, and therefore not observed by image analysis, until trypsin cleavage, at which point the polypeptides form a stable structure [12, 34]. It is possible that for astroviruses, the reverse takes place, and that fairly stable densities visible before cleavage become completely disordered at some positions. When the virus moves into the stomach of its host and is exposed to protease, the two arrangements of the spikes formed by AB and CC capsid protein dimers are likely to present different cleavage sites, giving rise to the two different C-terminal cleavage products, VP27 and VP25. The peripentonal spikes have dimer contributions that must extend nearly twice the distance of those at the icosahedral twofold axes, which may uncover sites for the additional trimming that generate VP25, and VP27 would be comprise the remaining visible spikes. The spikes located on the icosahedral twofolds are likely to be

VP27 homodimers, which have restricted movement just like the spikes in unprocessed capsid and thus can be visualized in a population-averaged reconstruction. There are several potential ways the conformational change could confer infectivity—by exposing the capsid surface, allowing some variable domains to be flexible, or exposure of binding sites.

Crystal Structure of the Human Astrovirus Capsid Spike

Structure Overview

Among the four small, nonenveloped positive-sense RNA viruses (i.e., picornavirus, calicivirus, astrovirus, and hepevirus) that infect animals, astrovirus is the only one whose atomic structure is not yet fully resolved. Based on primary sequence analysis, it was hypothesized that VP34 builds up the continuous capsid shell, while VP27 and VP25 form the dimeric projections seen in the cryo-EM reconstruction image [14, 23]. In an attempt to express the projection/spike domain, Dong et al. made five constructs containing residues 394–646, 501–658, 415–628, 501–628, and 415–646, respectively [13]. Of these five constructs, only $P2^{415–646}$ produced soluble proteins in *E. coli*, presumably because it encompasses an intact spike domain, and its two ends both terminate in flexible domain linkers. Considering that the N-terminal ends of VP27 and VP25 start from Gln394 and Ser424, respectively [27], $P2^{415–646}$ is a little longer than VP25 but a bit shorter than VP27 at the N-terminus. $P2^{415–646}$ exists as a dimer in solution as judged by gel filtration chromatography [13].

Purified $P2^{415–646}$ dimer was crystallized in two different crystal forms of space group $P3_221$ and $I2_13$, respectively. The $I2_13$ crystals, grown at pH 4.6, diffracted to only 2.8 Å possibly due to its high solvent content (82.8%). The $P3_221$ crystals, which were grown at pH 9.5, diffracted far better to 1.8 Å resolution (solvent content = 41.0%). Both structures have been determined, but it was found that the two structures were essentially the same with a RMSD of only 0.548 Å for all Ca atoms, despite the fact that these two crystals were grown at very different pHs and had different crystal packing interactions. Amino acid residues 430–645 of $P2^{415–646}$ were modeled in the crystal structure. As in solution, $P2^{415–646}$ formed dimers in both crystal forms around the twofold crystal symmetry axes. As discussed below, the crystal structure of the astrovirus $P2^{415–646}$ spike provides a first glimpse into the astrovirus capsid structure at the atomic level (PDB ID: 3QSQ).

The astrovirus surface spike has an exclusively β-stranded structure. Each $P2^{415–646}$ molecule contains a core, six-stranded β-barrel (β1, β4, β6, β7, β10, and β11) with a tightly packed hydrophobic core (Fig. 3.3). Another three-stranded β-sheet, consisting of β2, β3, and β5, is packed against the outside of the β-barrel away from the dimer interface. Therefore, the structure of $P2^{415–646}$ can be formalized as a triple-layered β-sandwich, in which the three orthogonal β-sheets are connected by two highly bent β-strands (β4 and β6). Two long loops, L1 and L4 (L1: residues

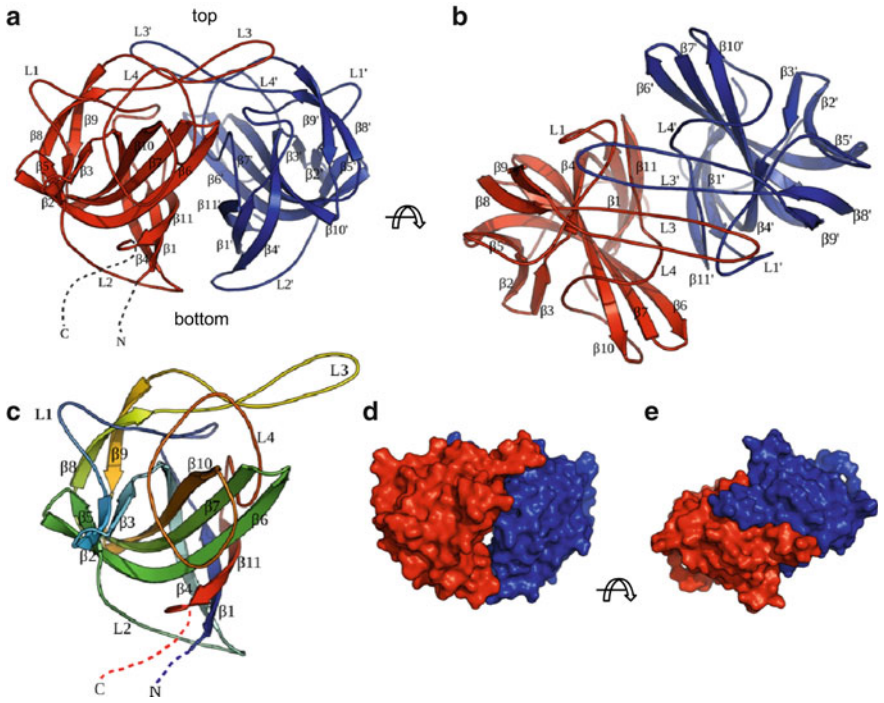

Fig. 3.3 Crystal structure of astrovirus $P2^{415–646}$. (**a**) Side view of the $P2^{415–646}$ dimer. The two subunits are colored in *red* and *blue*, respectively. The N- and C-termini of the *red* subunit are highlighted by *dotted lines*. The viral capsid lies at the bottom the dimeric spike. (**b**) Top view of the $P2^{415–646}$ dimer. (**c**) Structure of one $P2^{415–646}$ molecule. The polypeptide is *rainbow*-colored with the N-terminus in *blue* and the C-terminus in *red*. (**d**, **e**) Surface representations of the $P2^{415–646}$ dimer for comparison with the cryo-EM reconstruction image. Molecules in (**d**) and (**e**) are viewed from the same orientation as in (**a**) and (**b**), respectively

439–466; L4: residues 597–632), span across the top of the β barrel, while a shorter loop, L2 (residues 491–506), runs across the bottom of the barrel (Fig. 3.3). A β-hairpin motif, consisting of β8, β9, and L3 (residues 554–574), sits on top of L1 and L4, and may help to reinforce the structural conformation of these two long loops.

The $P2^{415–646}$ dimer appears to be very stable. The buried surface area at the dimer interface is ~3,490 $Å^2$, which counts for 20.2% of the solvent-accessible area of the complex. Therefore, the dimer interaction mediated by the projection domain is likely to play an important role in stabilizing the viral capsid. Because the astrovirus projection domain is connected to the capsid shell at its N-terminus, the thinner end of the projection dimer (where both the N- and C-termini are located) should directly abut the capsid outer surface, whereas the opposite, flat end would point away from the capsid (Fig. 3.3a). When visually compared to a cryo-EM reconstruction image of an astrovirus (Fig. 3.1), the overall morphology of the $P2^{415–646}$ dimer matches that of the dimeric surface spikes.

The projection dimer interface is primarily mediated by loops L3 and L4 and strands β1 and β11 (Fig. 3.3). In particular, two L3 loops from the two subunits of a dimer reach over to the opposing subunit like two crossing arms, making extensive interactions along their entire length. Sitting on top of the projection domain dimer, L3 also makes close contacts with the L1 and L4 loops from both subunits. A total of 51 residues from each subunit are observed at the dimer interface, among which 29 of the residues are highly conserved. Among the nonconserved interface residues, 12 residues interact with the other subunit via backbone atoms. Except the protruding L3 loop, residues at the dimer interface create a nearly flat surface in between the two molecules that is mostly hydrophobic in nature. The protruding L3 loop, however, contains mostly polar/charged residues and thus predominantly makes hydrogen bonds at the dimer interface. For example, 20 of the 26 intermolecular hydrogen bonds involve residues from the protruding L3 loop. Intermolecular hydrogen bonds are mediated by either main-chain atoms or side-chain atoms of highly conserved residues (i.e., Gln450, Tyr555, Arg558, and Arg572). The highly conserved nature of these inter-subunit hydrogen bonds indicates an important functional role and suggests that they may determine the specificity of dimer interaction.

Putative Receptor Binding Site(s)

Little is known about astrovirus receptor(s) and virus-receptor binding, although previous antigenicity and sequence analysis implicate the P2 domain in cell attachment and receptor binding [5, 23, 33]. Aligning the human astrovirus $P2^{415–646}$ domain of serotype 8 to that of the other seven serotypes gave pairwise percentage identities ranging from 47% to 57%. It is likely that all human astroviruses use a common receptor for cell internalization, even different astrovirus serotypes might use different cell surface molecules for initial attachment [9]. By calculating the amino acid conservation scores for the P2 domain using the sequences of all eight human astrovirus serotypes, several conserved regions were observed [13]. The most interesting one, which was called the P-site, is located in a shallow groove on top of the P2 spike (Fig. 3.4). This region has a large number of conserved residues (Lys455, Ser 554, Thr575, and Glu610 strictly conserved in all eight serotypes), all with their side chains exposed to solvent. The location of this site also makes it highly accessible to cell receptors without steric hindrance. Another conserved surface patch, the S-site, is located on the side of the P2 spike. This site contains three strictly conserved aromatic residues (Tyr475, Trp604, and Trp606), but two of these (Trp604 and Trp606) have their side chains tucked inside and thus may have been retained for the folding of the protein rather than for ligand recognition. Interestingly, an exposed β-turn (Thr529–Asn530–Asn531–Arg532) between strands β6 and β7 is also well conserved. This β-turn and its adjacent regions have the highest B-factors among the entire structure.

A close inspection of the P-site shows that it is largely hydrophilic in nature, and is abundant in hydrogen bond donors and acceptors (Fig. 3.4d). In addition to the strictly conserved residues Lys455, Ser 554, Thr575, and Glu610, other residues

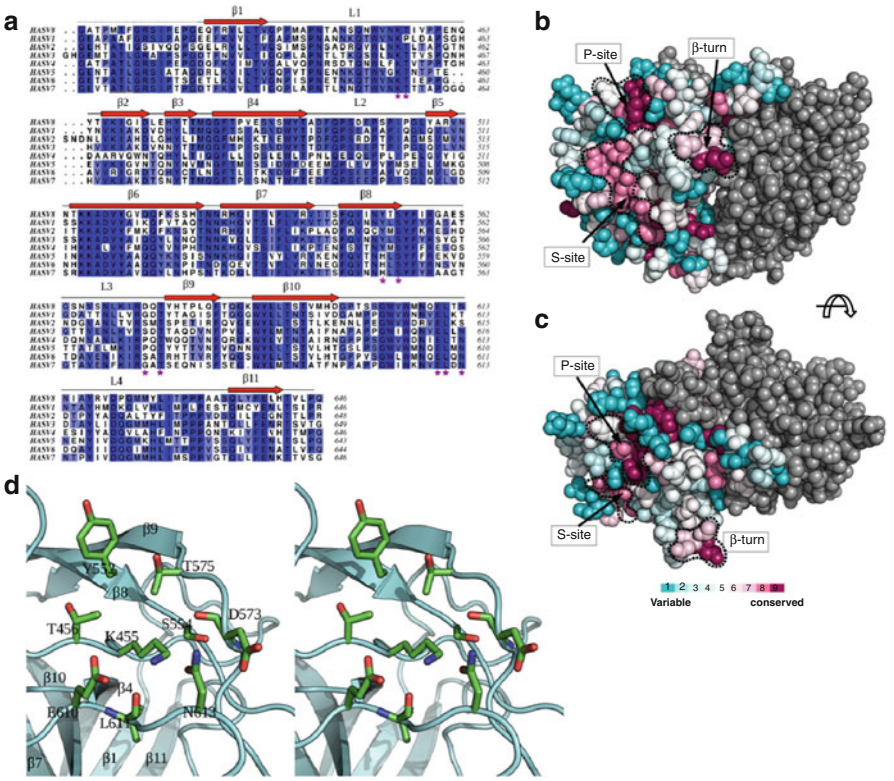

Fig. 3.4 Conserved surfaces on the human astrovirus projection domain. (**a**) Multiple sequence alignment of the projection domains (residues 415–646 for serotype 8) from human astroviruses of the eight different serotypes. The level of conservation is color-coded by different shades of *white* (least conserved) and *blue* (strictly conserved). The multiple sequence alignment was computed using Aline [8]. (**b, c**) Astrovirus projection domain, viewed from the side and the top, respectively. Residues in one subunit are colored according to their levels of conservation. The other subunit is colored in *gray*. Three highly conserved sites are circled by *dotted black curves*. (**d**) Stereo view of the putative receptor-binding site (P-site)

found in the vicinity are Thr456 (Pro in serotype 1), Tyr552 (His in serotypes 5, 6, and 7), and Asn613 (Thr in serotype 1, Ser in serotype 2), together with main-chain carbonyls from Asp573 and Leu611. These residues mainly come from the three loops L1, L3, and L4 (Fig. 3.4d). The polar residue composition, especially the positively/negatively charged pair Lys455 and Glu610, is also observed in the receptor binding sites of several viruses that use oligosaccharide moieties as receptors (e.g., influenza A virus and norovirus) [10, 17], suggesting that human astroviruses are likely to use a polysaccharide molecule as a cell receptor. Further molecular modeling and docking indicate that the P-site may be able to accommodate a di- or trisaccharide moiety. Consistent with structural analysis, the addition of heparin, heparan sulfate, or dextran sulfate was able to partially block HAstV-8 infectivity in a consistent manner [13]. Among the three polysaccharides used, heparin had the most pronounced effect.

Structure Comparison with HEV

Using the structure of the astrovirus capsid protein (CP) projection domain as a reference, a DALI search against all PDB entries indicated that the closest structural homolog is the projection domain of the HEV CP (z-score = 3.9). The structure of each HEV CP contains three domains, S, P1, and P2. The S domain adopts a typical jelly-roll β-barrel fold commonly found in other positive-stranded RNA viruses. The P1 domain has a structure of six-stranded β-barrel decorated with four short α-helices [19, 37]. The HEV CP P2 domain makes up the dimeric spikes on the viral surface and is responsible for receptor binding and antigenicity determination. The structural alignment generated by the DALI server shows that the HEV capsid P2 domain [19, 25, 37] has a core six-stranded β-barrel structure similar that of the astrovirus $P2^{415–646}$ (Fig. 3.5). Indeed, the matching β-strands in these two β-barrel structures (β1, β4, β6, β7, β10, and β11 for astrovirus, and β2, β3, β7, β9, β10, and β11 for HEV) have very similar spatial positions, although their lengths may be different (Fig. 3.5b). The strand β5 of astrovirus $P2^{415–646}$, although not part of the core six-stranded β-barrel, is also related to the strand β6 of HEV P2. In both structures, this β-strand is packed on the outer surface of the core β-barrel.

In addition to the similar structural folds, several major structural differences were also observed between the two viral spikes (Fig. 3.5b). First, the loop L2 that runs across the bottom of the core β-barrel of astrovirus $P2^{415–646}$ is replaced by two β-strands (β4 and β5) and a one-turn α-helix in HEV. Second, the long β-hairpin β6 and β7 in astrovirus $P2^{415–646}$ is replaced with two short strands (β7 and β9) with a loop in between that contains an extra β-strand (β8) in HEV. Third, there are two extra β-hairpin insertions in astrovirus, one between β1 and β4 and the other between β7 and β10. Fourth, β6, β7, and β10 in astrovirus CP are much longer than their counterparts in HEV P2. In astrovirus, these three β-strands form a small protruding knob on each side of the surface spikes. Because of these insertions and also because of its longer β-strands, the astrovirus P2 domain is substantially larger in size compared to the HEV P2 (232 vs. 143 residues), which could account for the extended radius of the HAstV spikes (Fig. 3.2).

The P2 dimer interaction is mediated by similar structural elements in both astrovirus and HEV, as structure superposition based on P2 monomers also placed the P2 dimers on top of each other, with the two dimer axes differing by only ~14°. Therefore, the inclination angle of the β-barrel with respect to the icosahedral twofold symmetry axis must be similar in astrovirus and HEV, but the azimuthal rotation of the dimeric spike about the icosahedral twofold axis is rather different in these two viruses. Like in the HEV CP, both N- and C-termini of the astrovirus P2 are located at the bottom of the spike close to the particle surface.

The detailed structural similarity between the astrovirus and HEV surface spikes suggests that the two viruses may have a much closer phylogenetic relationship than previously expected [22]. On the other hand, the caliciviruses, another group of small, nonenveloped, positive-sense RNA viruses, also have dimeric surface spikes. According to Dali, calicivirus CP is also related to astrovirus CP in its overall

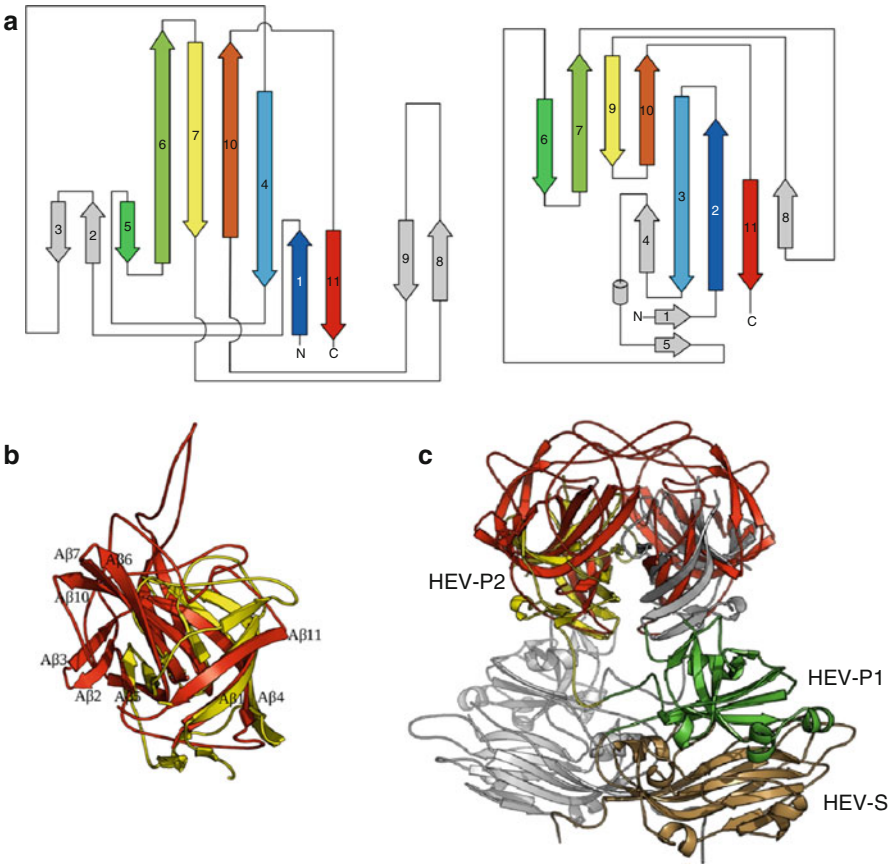

Fig. 3.5 Structural comparison of $P2^{415–646}$ with HEV capsid P2 domain. (**a**) Topology diagrams for the astrovirus $P2^{415–646}$ (*left*) and the HEV P2 domain (*right*). Related β-strands in the two structures are rendered in matching colors, while unrelated structural elements are shaded in *gray*. (**b**) Superposition of the astrovirus $P2^{415–646}$ (*red*) and the HEV P2 domain (*yellow*). Structure elements in astrovirus $P2^{415–646}$ are labeled as in Fig. 3.3c. with a letter "A". (**c**) Superposition of the astrovirus $P2^{415–646}$ dimer with the HEV CP dimer. The astrovirus $P2^{415–646}$ dimer is in *red*. For HEV CP, one subunit is in *gray*, and the other is colored by domains: S in *gold*; P1 in *green*; and P2 in *yellow*

structural fold, but the level of similarity is significantly less than that for HEV ($Z=1.7$ vs. $Z=3.9$). In addition, the domain organization of the calicivirus CP is different compared to that of astrovirus and HEV CP. While the three domains of the HEV CP are organized in a linear fashion, the polypeptide of the calicivirus CP winds through the three domains in the order of S → P1 → P2 → P1 [29]. Therefore, the surface spike of calicivirus is formed effectively by a polypeptide insertion in the P1 domain. In contrast, the spike of both HEV and astrovirus is situated on the CP polypeptide after the P1 domain (Fig. 3.6a). Based on these observations, astrovirus is likely to have more distant phylogenetic relationship to calicivirus compared to HEV, at least when the viral CPs are used for analysis.

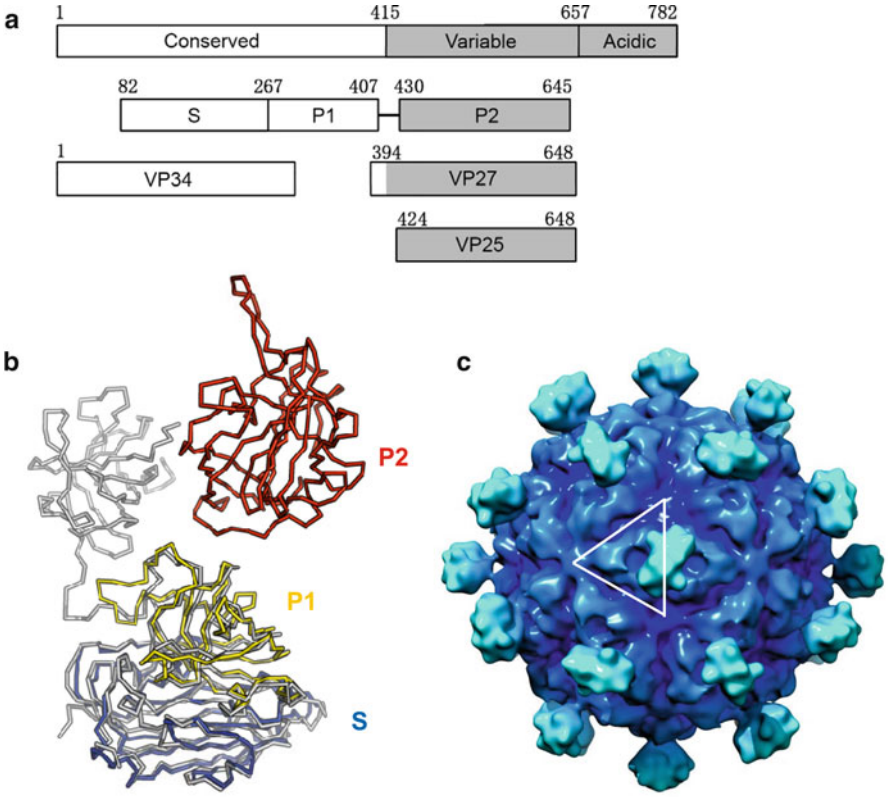

Fig. 3.6 Astrovirus capsid structure model. (**a**) Domain structure of astrovirus CP of serotype 8. Astrovirus CP has a conserved region, a variable region, and an acidic C-terminal region. The S and P1 domain of astrovirus can be roughly mapped to the conserved region. The P2 domain, however, is mapped to the variable region. The coverage of VP34, VP27, and VP24 is also indicated. (**b**) Astrovirus CP superimposed onto HEV CP. HEV CP is shown in *gray*, whereas the astrovirus CP is colored by domains with S in *blue*, P1 in *yellow*, and P2 in *red*. (**c**) $T=3$ astrovirus capsid model rendered at 15 Å resolution. The *white lines* highlight one asymmetric unit

Structure Modeling of the Human Astrovirus Capsid and Insights into Maturation

Based on the noted structural similarity between the astrovirus and HEV CPs, an attempt was made to create an atomic structure of the astrovirus CP by homology modeling [13]. Indeed, homology structure recognition based on the astrovirus CP sequence alone indicated that the HEV CP is the best structural homolog [13]. In particular, residues 82–266 and 267–407 of the human astrovirus (serotype 8) CP were matched to the S (aa118–313) and P1 (aa314–453) domains of the HEV CP, respectively. Based on this alignment, a homology model of the human astrovirus CP was created with the HEV CP structure as a template [13] (Fig. 3.6). By aligning the S and P1 domains of the astrovirus CP model onto the three quasi-equivalent

HEV CP molecules from the $T=3$ EM model [36], a $T=3$ astrovirus capsid structure was generated [13]. The astrovirus capsid model closely resembles the cryo-EM reconstruction of astrovirus in terms of particle diameter and the capsid landscape variation. The close resemblance of the two structures further confirms the notion that astroviruses are phylogenetically related to HEV.

During maturation, the immature viral particles are processed by extracellular proteases to produce smaller peptides species that are observed in infectious particles. VP34 is derived from the conserved region, while VP27 and VP25 are from the hypervariable region of the astrovirus CP [26] (Fig. 3.6). The N-terminal end of VP27 and VP25 starts from Gln394 and Ser424, respectively [27]. Based on the astrovirus capsid model, the N-terminal cleavage site of VP27 (Gln394) was mapped to the unstructured loop right before the last β-strand in the P1 domain. Therefore, VP27 should be able to remain associated with the viral particle through main-chain H-bonds mediated by this last β-strand in the P1 domain. The N-terminal cleavage site of VP25, however, is located in an extended linker (the P1–P2 linker) between the P1 and the P2 domain. Therefore, VP25 may need to dimerize with a VP27 to remain stably bound to the viral particle.

While the astrovirus CP is larger than HEV CP by ~120aa, there may be additional similarities in their capsid assembly processes. HastV-8 has been shown to have caspase processing at the $TYVD_{657}$ sequence [3], which removes the last ~125 amino acids for the VP90 conversion to VP70 and is required for viral release. Although this particular sequence is not conserved in all astroviruses, they all contain potential caspase cleavage sequences within amino acids 650–700 of their CP polypeptides. Because HEV assembly has only been achieved in vitro with the deletion of ~50 C-terminal amino acids, HEV may have a similar cleavage requirement for particle release. As very little is known about HEV infectivity, it is tempting to speculate that those viruses may also require extracellular protease cleavage for activation, which may be similarly asymmetric to that observed for astrovirus.

Avian Astrovirus Capsid Spike Domain Structure

Recently, the crystal structure of the turkey astrovirus 2 capsid spike domain was determined at 1.5 Å resolution and publicly released by the Protein Data Bank (PDB ID: 3TS3). Although the corresponding publication is not yet published, visual inspection of the structure reveals significant structural divergence compared to the HAstV-8 capsid spike structure (PDB ID: 3QSQ) [13]. Whereas the human spike domain is composed entirely of β-strands and forms a bowl-shaped dimeric spike, the avian spike domain contains several helices as well as β-strands and forms a heart-shaped dimeric spike (Fig. 3.7). In fact, a structural homology search identifies the HEV spike domain, not the human astrovirus spike domain, as the closest structural homolog of the avian astrovirus spike domain. The low sequence homology and low structural conservation between the mammalian and avian astrovirus spike domains suggest that mammalian and avian astroviruses may have significant mechanistic differences in host cell entry.

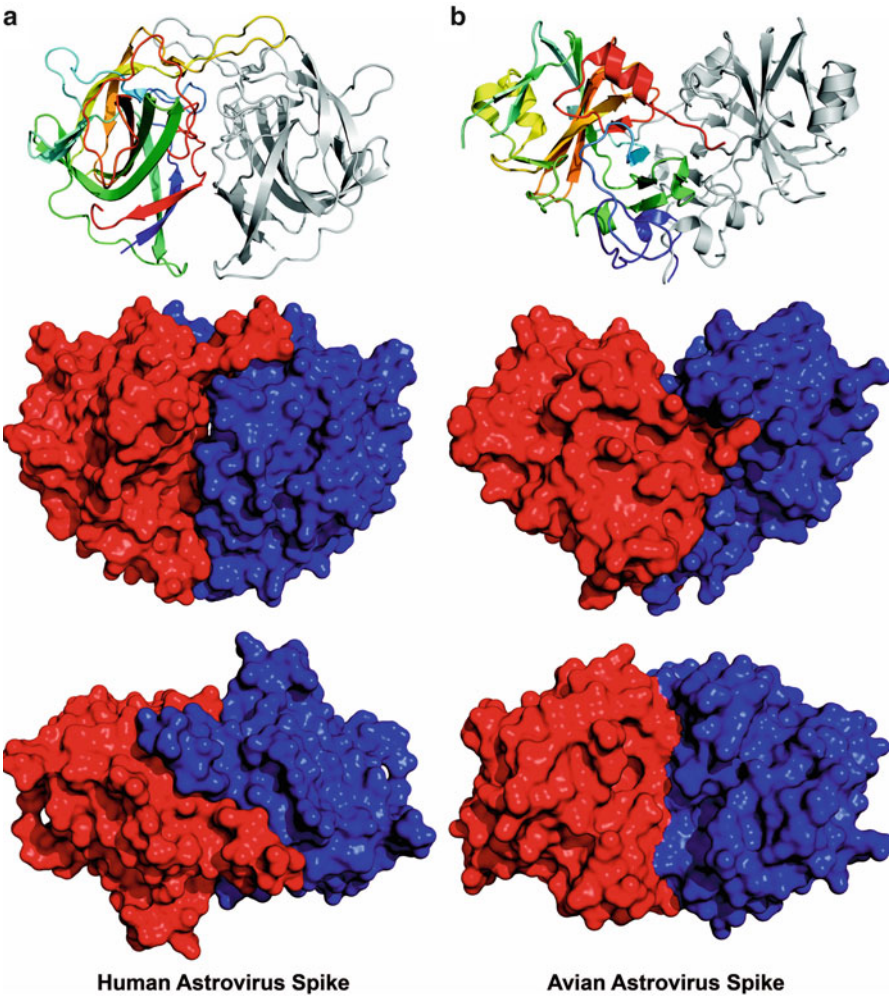

Fig. 3.7 Comparison of human and avian astrovirus capsid spike structures. (**a**) Dimeric structure of the human astrovirus 8 capsid spike shown as cartoon (*top*) or surface (*bottom*) representation. (PDB entry 3QSQ). (**b**) Dimeric structure of the turkey astrovirus 2 spike shown as cartoon (*top*) or surface (*bottom*) representation (PDB entry 3TS3). In the cartoon representations, one protomer is colored in a *blue*-to-*red rainbow* from N- to C-termini whereas the other protomer is colored *gray*. In the surface representations, one protomer is colored *red* and the other is colored *blue*

Human Astrovirus Protease Structure

The astrovirus genome open reading frames ORF1a and ORF1b encode the nonstructural proteins nsp1a and nsp1b, respectively, which are involved in viral replication. The nonstructural proteins are synthesized as polyprotein precursors that become proteolytically processed. Although the exact processing sites remain

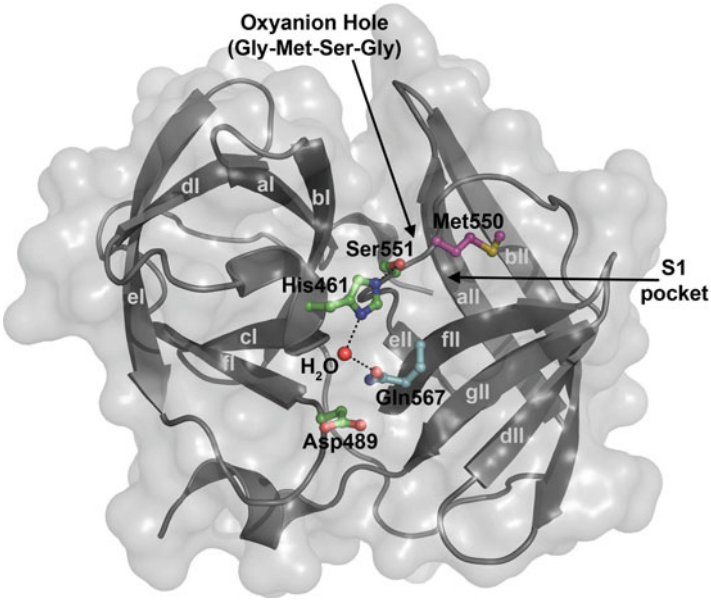

Fig. 3.8 Structure of the astrovirus protease. Human astrovirus 1 protease structure (ORF1a polyprotein residues 432–587) shown as cartoon representation overlaid with a semitransparent surface representation. The catalytic triad (His461–Asp489–Ser551) side chains are shown as *green ball*-and-stick models. The Gln567 side chain that hydrogen bonds to a water molecule in the active site is shown as *cyan ball*-and-*stick* model. The water molecule is shown as a *red sphere*. The Met550 side chain of the oxyanion hole (Gly549–Met550–Ser551–Gly 552) is shown as *magenta ball*-and-*stick* model. β-Strands are labeled in *light gray*. Locations of the oxyanion hole and positively charged S1 substrate binding pocket are described with *arrows*

undefined, a 3C-like serine protease motif was found to be encoded at the 3′-end of ORF1a and is thought to play an important role in the processing events [20]. Further studies show that mutation of the astrovirus protease motif resulted in deficient or altered processing of nsp1a translated in vitro [20] or in BHK cells [18].

The human astrovirus 1 protease crystal structure was determined at 2.0 Å resolution (PDB entry 2W5E) [35]. The 17 kDa protein is a member of the trypsin-like serine protease family and contains two β-barrel domains connected by a long interdomain loop (Fig. 3.8). The active site contains the catalytic triad His461, Asp489, and Ser551 that was predicted by Kiang et al. (1992), and these residues are strictly conserved in all astroviruses [20]. In addition, the Gly–X–Ser–Gly motif conserved in all astroviruses at residues 549–552 that forms the oxyanion hole is in the correct structural conformation for stabilization of the intermediate during catalysis.

However, the astrovirus protease has several structural features that distinguish it from other viral proteases [35]. First, two loops near the active site, loop eI-fI and loop cII-dII, are shorter than most viral proteases. The result of this is a more open and solvent-exposed active site. Second, the active site residue Asp489 is oriented away from His461 and does not form the hydrogen bond normally seen in trypsin-like

proteases. Instead, a water molecule forms the hydrogen bond. The authors note that Asp489 likely undergoes a conformational change during substrate binding, as seen in hepatitis A virus protease [6, 38], although the authors do not rule out the possibility of an atypical catalytic mechanism. Last, the astrovirus protease substrate binding pocket S1 is more shallow than other viral proteases and has a positive electrostatic potential. Consistent with this structural observation, hydrolytic activity assays show that the astrovirus protease S1 pocket has specificity for negatively charged Asp and Glu residues. Furthermore, the hydrolytic activity is dependent on a functional astrovirus protease active site because the Ser551Ala mutant was inactive.

Taken together, the astrovirus protease structure reveals a trypsin-like serine protease with a specificity for cleavage at negatively charged residues Asp and Glu. Several distinctive structural features of the astrovirus protease may allow for further specificity during cleavage of the nonstructural polyprotein processing events.

Summary

Astrovirus structure and assembly are particularly intriguing in several aspects. First, a ~125aa C-terminal domain of the viral CP is removed by host caspase following assembly. The exact function of this C-terminal domain is not known but it may facilitate the sorting/trafficking of the viral CP to promote assembly. Additionally, extensive extracellular protease processing of the viral capsid is required for virus infectivity and maturation, which is rather unusual for small, positive-sense RNA viruses. In this chapter, we have primarily focused our discussion on the cryo-EM reconstructions of the immature and mature astrovirus, the crystal structures of the human astrovirus and turkey astrovirus capsid spikes, and the crystal structure of the human astrovirus protease. Altogether, these new findings have led a better understanding of many facets of the astrovirus replication, including astrovirus assembly, virus release, maturation, receptor binding, antibody neutralization, and nonstructural polyprotein processing. Of course, these findings would not be possible without many years of pioneering virological work that is not fully discussed here. We also realize that the ultimate understanding of the astrovirus assembly and maturation will require further characterization of the viral capsid/capsid proteins by structural as well as biochemical, genetic, and cell biology approaches.

Although there is a clear structural homology between the astrovirus and HEV capsid proteins, the replication mechanism of the astrovirus appears to differ from that of HEV. Unlike HEV, which is known to encode both methyltransferase and guanylyltransferase activities within the 5′ half of its ORF1, no capping enzymes have been identified in the astrovirus nonstructural proteins. Only three enzymes, the viral protease, a viral polymerase, and a putative helicase, have been found in the astrovirus genome. Interestingly, phylogenetic analysis suggests that the astrovirus genome might encode a VPg with sequence similar to the VPgs from *Caliciviridae* and *Potyviridae* [1]. We expect that further studies on astrovirus and

other small, nonenveloped RNA viruses, both on their capsid and their replication, would help to establish their phylogenetic relationships, which, in turn, would greatly facilitate further structural and functional characterization of these viruses.

Acknowledgments The authors are supported by the Welch Foundation (C-1565 to YJT) and the National Institutes of Health (R01 GM066087 to MY).

References

1. Al-Mutairy B, Walter JE, Pothen A, Mitchell DK. Genome prediction of putative genome-linked viral protein (VPg) of astroviruses. Virus Genes. 2005;31:21–30.
2. Appleton H, Higgins PG. Letter: Viruses and gastroenteritis in infants. Lancet. 1975;1:1297.
3. Banos-Lara Mdel R, Méndez E. 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. 2010;401:322–32.
4. Bass DM, Qiu S. Proteolytic processing of the astrovirus capsid. J Virol. 2000;74:1810–4.
5. Bass DM, Upadhyayula U. Characterization of human serotype 1 astrovirus-neutralizing epitopes. J Virol. 1997;71:8666–71.
6. Bergmann EM, Mosimann SC, Chernaia MM, Malcolm BA, James MN. The refined crystal structure of the 3C gene product from hepatitis A virus: specific proteinase activity and RNA recognition. J Virol. 1997;71:2436–48.
7. Bodkin DK, Nibert ML, Fields BN. Proteolytic digestion of reovirus in the intestinal lumens of neonatal mice. J Virol. 1989;63:4676–81.
8. Bond CS, Schuttelkopf AW. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. Acta Crystallogr D Biol Crystallogr. 2009;65:510–2.
9. Brinker JP, Blacklow NR, Herrmann JE. Human astrovirus isolation and propagation in multiple cell lines. Arch Virol. 2000;145:1847–56.
10. Cao S, Lou Z, Tan M, Chen Y, Liu Y, Zhang Z, et al. Structural basis for the recognition of blood group trisaccharides by norovirus. J Virol. 2007;81:5949–57.
11. Caul EO, Appleton H. The electron microscopical and physical characteristics of small round human fecal viruses: an interim scheme for classification. J Med Virol. 1982;9:257–65.
12. Crawford SE, Mukherjee SK, Estes MK, Lawton JA, Shaw AL, Ramig RF, et al. Trypsin cleavage stabilizes the rotavirus VP4 spike. J Virol. 2001;75:6052–61.
13. Dong J, Dong L, Méndez E, Tao Y. Crystal structure of the human astrovirus capsid spike. Proc Natl Acad Sci USA. 2011;108:12681–6.
14. Dryden KA, Tihova M, Nowotny N, Matsui SM, Méndez E, Yeager M. Immature and mature human astrovirus: structure, conformational changes, and similarities to Hepatitis E Virus. J Mol Biol. 2012; Epub ahead of print June 25.
15. Du L, Kao RY, Zhou Y, He Y, Zhao G, Wong C, et al. Cleavage of spike protein of SARS coronavirus by protease factor Xa is associated with viral infectivity. Biochem Biophys Res Commun. 2007;359:174–9.
16. Estes MK, Graham DY, Mason BB. Proteolytic enhancement of rotavirus infectivity: molecular mechanisms. J Virol. 1981;39:879–88.
17. Gamblin SJ, Haire LF, Russell RJ, Stevens DJ, Xiao B, Ha Y, et al. The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science. 2004;303:1838–42.
18. Geigenmüller U, Ginzton NH, Matsui SM. Studies on intracellular processing of the capsid protein of human astrovirus serotype 1 in infected cells. J Gen Virol. 2002;83:1691–5.
19. Guu TSY, Liu Z, Ye Q, Mata DA, Li K, Yin C, et al. Structure of the hepatitis E virus-like particle suggests mechanisms for virus assembly and receptor binding. Proc Natl Acad Sci USA. 2009;106:12992–7.

20. Kiang D, Matsui SM. Proteolytic processing of a human astrovirus nonstructural protein. J Gen Virol. 2002;83:25–34.
21. Klenk HD, Rott R, Orlich M, Blodorn J. Activation of influenza A viruses by trypsin treatment. Virology. 1975;68:426–39.
22. Koci MD, Seal BS, Schultz-Cherry S. Molecular characterization of an avian astrovirus. J Virol. 2000;74:6173–7.
23. Krishna NK. Identification of structural domains involved in astrovirus capsid biology. Viral Immunol. 2005;18:17–26.
24. Lazarowitz SG, Choppin PW. Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. Virology. 1975;68:440–54.
25. Li S, Tang X, Seetharaman J, Yang C, Gu Y, Zhang J, et al. Dimerization of hepatitis E virus capsid protein E2s domain is essential for virus-host interaction. PLoS Pathog. 2009;5:e1000537.
26. Méndez E, Aguirre-Crespo G, Zavala G, Arias CF. Association of the astrovirus structural protein VP90 with membranes plays a role in virus morphogenesis. J Virol. 2007;81:10649–58.
27. Méndez E, Fernández-Luna T, López S, Méndez-Toss M, Arias CF. Proteolytic processing of a serotype 8 human astrovirus ORF2 polyprotein. J Virol. 2002;76:7996–8002.
28. Méndez E, Salas-Ocampo E, Arias CF. Caspases mediate processing of the capsid precursor and cell release of human astroviruses. J Virol. 2004;78:8601–8.
29. Prasad BV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK. X-ray crystallographic structure of the Norwalk virus capsid. Science. 1999;286:287–90.
30. Risco C, Carrascosa JL, Pedregosa AM, Humphrey CD, Sánchez-Fauquier A. Ultrastructure of human astrovirus serotype 2. J Gen Virol. 1995;76(Pt 8):2075–80.
31. Royuela E, Sánchez-Fauquier A. Molecular cloning, expression and first antigenic characterization of human astrovirus VP26 structural protein and a C-terminal deleted form. Comp Immunol Microbiol Infect Dis. 2010;33:1–14.
32. Salminen A, Wahlberg JM, Lobigs M, Liljestrom P, Garoff H. Membrane fusion process of Semliki Forest virus. II: cleavage-dependent reorganization of the spike protein complex controls virus entry. J Cell Biol. 1992;116:349–57.
33. Sanchez-Fauquier A, Carrascosa AL, Carrascosa JL, Otero A, Glass RI, Lopez JA, et al. Characterization of a human astrovirus serotype 2 structural protein (VP26) that contains an epitope involved in virus neutralization. Virology. 1994;201:312–20.
34. Settembre EC, Chen JZ, Dormitzer PR, Grigorieff N, Harrison SC. Atomic model of an infectious rotavirus particle. EMBO J. 2011;30:408–16.
35. Speroni S, Rohayem J, Nenci S, Bonivento D, Robel I, Barthel J, et al. Structural and biochemical analysis of human pathogenic astrovirus serine protease at 2.0 A resolution. J Mol Biol. 2009;387:1137–52.
36. Xing L, Li T-C, Mayazaki N, Simon MN, Wall JS, Moore M, et al. Structure of hepatitis E virion-sized particle reveals an RNA-dependent viral assembly pathway. J Biol Chem. 2010;285:33175–83.
37. Yamashita T, Mori Y, Miyazaki N, Cheng RH, Yoshimura M, Unno H, et al. Biological and immunological characteristics of hepatitis E virus-like particles based on the crystal structure. Proc Natl Acad Sci USA. 2009;106:12986–91.
38. Yin J, Bergmann EM, Cherney MM, Lall MS, Jain RP, Vederas JC, et al. Dual modes of modification of hepatitis A virus 3C protease by a serine-derived beta-lactone: selective crystallization and formation of a functional catalytic triad in the active site. J Mol Biol. 2005;354:854–71.