

Chapter 22

Astrovirus Replication: An Overview

Susana Guix[†], Albert Bosch^{,†} and Rosa M. Pintó[†]*

Human astroviruses are important pathogens that cause gastroenteritis worldwide. Significant progress has recently been made regarding the characterization of the RNA replication process, the apoptotic response induced in virus-infected cells, and the formation of virus-like particles. First, a relationship between astrovirus RNA replication sites and the endoplasmic reticulum-derived membranes has been suggested. In addition, a direct relationship between nonstructural proteins and the induction of apoptosis has been observed, and it has been demonstrated that apoptotic host cell death seems necessary for maturation of astrovirus particles. Finally, it has been predicted that the VP34 capsid protein contains the RNA binding domain at its N-terminus, responsible for packaging the viral genome, as well as an 8 β -barrel domain that may likely constitute the building subunit to form the T = 3 icosahedral capsid. The C-terminal half of the capsid polyprotein — highly variable between different astroviruses — is thought to form the receptor-interaction domain.

Background for Human Astroviruses

Infections of the gastrointestinal tract cause approximately two billion cases of diarrhea in children per year worldwide, with the majority of

*Corresponding author: Albert Bosch, Department of Microbiology, University of Barcelona, Avda Diagonal 645, 08028 Barcelona, Spain, E-mail: abosch@ub.edu.

[†]Enteric Virus Group, Department of Microbiology, University of Barcelona.

deaths occurring in developing countries.¹ Diarrhea is the third leading cause of mortality worldwide. Every year, an estimated three million pediatric deaths result from viral gastroenteritis and dehydration.² Along with rotavirus, calicivirus, and enteric adenovirus, human astroviruses (HAstV) are one of the most important causes of viral pediatric acute gastroenteritis.³⁻⁵ HAstV were first identified 30 years ago by Appleton and Higgins⁶ as the cause of a gastroenteritis outbreak in a maternity ward. Shortly thereafter, Madeley and Cosgrove⁷ proposed the name “astrovirus” after observing the virus’ star-like morphology by electron microscopy (EM) in stool samples from a pediatric diarrhea outbreak. Despite the years that have elapsed since HAstV were discovered, many aspects of their molecular biology and pathogenesis remain still to be elucidated.

Human astroviruses are non-enveloped icosahedral viruses that belong to the family *Astroviridae* with a plus-sense, single-stranded RNA genome. According to the International Committee on Taxonomy of Viruses, the *Astroviridae* family is divided into two genera: *Mamastrovirus*, which includes astroviruses that infect mammals and primarily causes gastroenteritis, and *Avastrovirus*, which includes astroviruses that infect avian species and may cause intestinal as well as extra-intestinal illness. Presently, HAstV are divided into eight serotypes (HAstV-1 to HAstV-8), with HAstV-1 being the most prevalent worldwide. HAstV-1 to 7 were initially identified according to the reactivity of the capsid proteins with reference type-specific rabbit antibodies, while HAstV-8 was fully characterized some years later.⁸⁻¹⁰ Over the last few years, the number of astrovirus sequences available via Genbank has rapidly increased and the complete genome for five HAstV isolates — including serotypes 1, 2, 3, and 8 — has been sequenced.^{9,11-14} Complete capsid sequences are available for HAstV-1 through 6 and HAstV-8, as well as for feline, porcine, ovine, mink, and turkey astroviruses and for the avian nephritis virus. HAstV can be isolated and propagated in human intestinal CaCo-2 continuous cell line in the presence of trypsin, which is involved in the capsid protein maturation.¹⁵ An infectious full-length cDNA clone is available for HAstV-1.¹⁶ BHK-21 cells can support HAstV-1 RNA

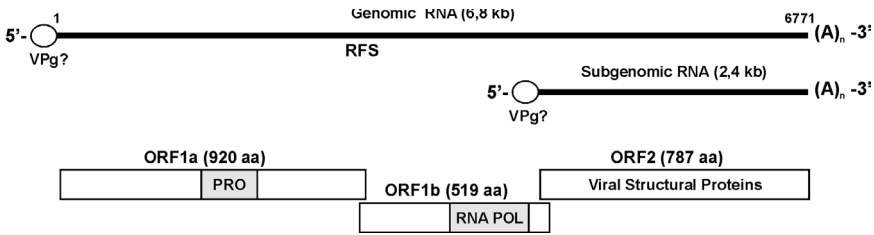


Fig. 1. Genomic organization of HAstV. Genomic and subgenomic RNA molecules are indicated in the upper panel, and ORFs with their main functional motifs are shown in rectangles. Ribosomal frame-shifting signal (RFS) between ORF1a and ORF1b is indicated. The presence of a VPg protein at the 5' end of RNA has not been confirmed. Nucleotide and amino acid positions are numbered according to HAstV-1 Oxford reference strain (accession no. L23513).

replication when transfected with *in vitro* transcribed RNA, and virions produced in these cells can infect CaCo-2 cell monolayers in trypsin-containing media.

Virions of HAstV contain a positive sense single-stranded 6.8 kb RNA genome, which is polyadenylated at the 3'-end and includes a 5'-untranslated region (UTR), three overlapping open reading frames (ORFs), and a 3'-UTR (Fig. 1). Although suggested, the presence of a covalently linked VPg protein at the 5'-end of the genome has not yet been biologically demonstrated.¹⁷ The fact that *in vitro* transcribed capped RNA is infectious does not rule out the possibility that the 5' end of genomic RNA found within particles could be linked to a VPg protein, and similar observations have been made with feline calicivirus.¹⁸ ORF1a and ORF1b are linked by a ribosomal frame-shifting signal (RFS) and code for the nonstructural proteins, including a serine protease and an RNA-dependent RNA polymerase, while ORF2 encodes the capsid precursor.¹¹

Although it is considered characteristic of positive-stranded RNA viruses with genomes larger than 6 kb to encode a helicase domain, there was a widespread belief for many years that astroviruses did not code for a helicase protein. However, two putative helicase conserved

motifs upstream from the protease coding region, which show sequence similarity to pestivirus RNA helicase sequences, have recently been mapped by computational tools.¹⁷

Upon infection, two large nonstructural proteins are translated from the genomic RNA: nsP1a and nsP1a1b. These proteins mature by proteolysis involving both the viral protease and cellular proteases, and give rise to the viral proteins implicated in the transcription of a full-length negative-strand RNA molecule (antigenomic RNA). It is believed that this RNA molecule serves as a template for the transcription of both the new full-length 6.8 kb-genomic and the ORF2-containing 2.4 kb-subgenomic RNA molecule.¹⁹ Translation of this subgenomic RNA is used to produce large amounts of structural proteins for the efficient assembly of the progeny viruses.¹⁵ Both astrovirus genome organization and replication/translation strategies have frequently been compared to those of alphaviruses and caliciviruses due to their similarities.

This review will focus on recent advances in the understanding of the molecular biology, structure, and replication cycle of HAstV. Current efforts are being undertaken to complete the characterization of their genomic organization, to study virion assembly and morphogenesis, and to understand the relationships between the virus and the host immune response. Major developments on these issues are expected in the near future.

Molecular Virology and Replication Cycle: New Insights

Structural Proteins

Considerable progress has been made in understanding the basics of HAstV genomic organization; however, capsid protein processing and assembly, as well as genome encapsidation, are not yet clear. Translation of the structural polyprotein from ORF2 (782 to 794 amino acids, depending on serotype) most likely occurs from the subgenomic RNA. The presence of a conserved RNA sequence upstream from the ORF2 initiation codon with a possible role in the regulation of the

capsid gene expression has been reported.²⁰ This region has also been identified as a hotspot for RNA recombination.²¹

Proteolytic Processing

As illustrated in Fig. 2, the polyprotein precursor would be cleaved intracellularly by cellular proteases and extracellularly by trypsin to produce three mature structural proteins of 32–34 kDa (VP34), 29–31 kDa (VP29), and 24–26 kDa (VP26), with molecular weights and trypsin cleavage sites varying between different serotypes.^{8,15,22–24} In studies based on HAstV-2, it was found that the 87 kDa polyprotein could be assembled into particles and then cleaved into proteins of 32, 29 and 26 kDa.^{24,25} Bass and Qiu suggested that assembly of HAstV-1 particles requires an intracellular cleavage of the first 70 N-terminal amino acids of the 87 kDa precursor protein.²² The resultant 79 kDa protein would assemble into the viral capsid within infected cells. In the absence of trypsin, these particles are virtually noninfectious, but they can become infectious when trypsin cleaves the 79 kDa protein into the final structural proteins VP34, VP29 and VP26. This intracellular cleavage at the N-terminus was not observed in other studies with HAstV-1²⁶ and HAstV-8⁸ infected cells, but it has recently been demonstrated for HAstV-1 that the 70 amino acids at the N-terminus of the polyprotein are not essential to assemble into virus-like particles.²⁷

An extensive study using HAstV-8 demonstrated that the 90 kDa polyprotein (VP90) is first cleaved intracellularly at its C-terminus, giving rise to a 70 kDa protein (VP70) that is mainly found into purified particles.⁸ The VP70-containing virus is minimally infectious and requires trypsin to enhance its infectivity. Following trypsin activation, VP70 is processed into the three predominant smaller mature products. In a later study, the same authors observed that indeed the capsid precursor VP90 is able to assemble into virions, as it had been shown to do in other studies for HAstV-1²⁶ and HAstV-2.^{24,25} However, these resulting particles were relatively unstable and would disassemble during cesium chloride purification, indicating that the cleavage that yields VP70 structurally stabilizes the viral particles.²³

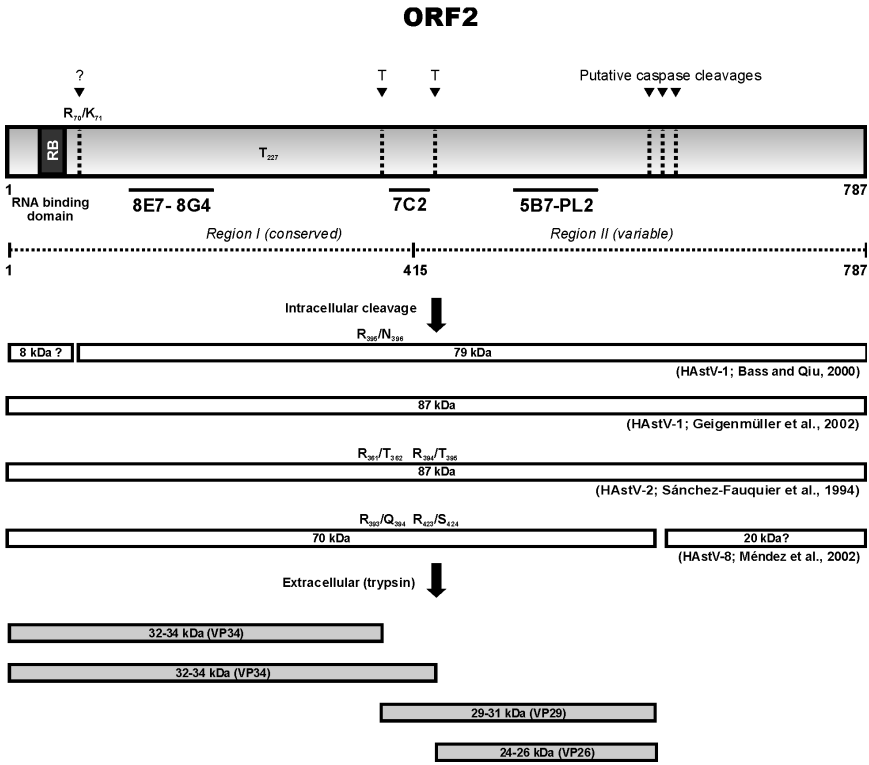


Fig. 2. Schematic depiction of the ORF2 polyprotein of HAstV, and proteolytic processing pathways proposed by different authors using different serotypes (Geigenmüller *et al.*, 2002).^{8,22,24} Putative RNA binding domain (RB) at the N-terminus of the polyprotein is indicated. Inverted arrowheads denote putative cleavage sites dependent upon caspases, other unknown cellular proteases (?), and trypsin (T). Underlined regions correspond to the main viral epitopes identified by different authors,^{24,37,38} being the epitopes defined by the mAb 8E7 and 8G4 shared by all serotypes. The dotted line indicates the two domains of the protein sequence (amino acids 1–415, well conserved; amino acids 416–end, variable). White boxes represent intermediate products found within infected cells, while grey boxes represent the final products present in infectious particles. See text for details. Amino acid positions are numbered according to HAstV-1 Oxford reference strain (accession no. L23513).

Capsid Protein Maturation and Apoptosis

The fact that HAsV infection induces apoptosis on CaCo-2 cells and that apoptotic death of host cells seems necessary for efficient human astrovirus replication and particle maturation has been demonstrated by different authors.^{23,28} Guix *et al.* observed that in the presence of a caspase 8-specific inhibitor, there was a significant reduction in the infectivity of the virus progeny, while the titer of physical virions in cell supernatants estimated by ELISA and RT-PCR was not affected.²⁸ Méndez *et al.* demonstrated that VP90-VP70 processing is mediated by caspases, and also observed that the release of infectious viruses to the cell supernatant was reduced in the presence of a pancaspase inhibitor and increased by the TNF-related apoptosis-inducing ligand (TRAIL).²³ However, inhibition of apoptosis also resulted in a decrease in the total amount of viral protein detected in cell supernatants, suggesting that processing of VP90 and/or the apoptotic cell response may also affect viral release.²³ Differences between the conditions of infections and the techniques used to measure the amount of released viruses could partially explain these discrepancies. For several viruses it has been observed that some of their proteins, both structural and nonstructural, undergo caspase-mediated proteolysis within the host cell,²⁹⁻³⁴ and in some cases such as human influenza virus or Aleutian mink disease parvovirus, these cleavages are relevant for virus replication and morphogenesis.^{30,34}

Structural Studies

Very little is known about the structural features of the HAsV capsid, and most of our current knowledge comes from structural predictions and comparison to other related viral families (for a review see Krishna, 2005).³⁵ Based on the homology between different astroviruses, the ORF2 sequence is divided into two regions. At the N-terminus, region I spans amino acid residues 1 to 415 and is highly conserved among all members of the *Astroviridae* family, while region II at the C-terminus is extremely variable among different strains

(Fig. 2). Wang *et al.* further subdivided region II into three regions with different degrees of similarity.³⁶ Consistent with the capsid polyprotein domain organization, studies using antibodies revealed that VP34 would express conserved epitopes shared by all serotypes, while VP29 and VP26 contained the serotype-specific neutralizing epitopes.^{24,37,38} Thus, it is reasonable to hypothesize that the hyper-variable C-terminal region would be located on the surface of the viral particle contributing to the strain-specific tropism of the virus, while the conserved region would constitute the “assembly domain,” a building block for capsid formation and encapsidation of the viral genome.^{27,35} Interestingly, during the construction of the HAstV-1 infectious clone it was found that amino acid residue Thr₂₂₇, within region I of the capsid polyprotein, was essential for proper assembly of viral capsids.¹⁹

Finally, within the first 70 residues of the ORF2 polyprotein, region I includes a well-conserved domain rich in basic amino acids, which has been associated with the viral RNA packaging process due to its potential RNA-binding properties and its similarities to other well-documented icosahedral RNA viruses, such as Sindbis virus and some plant viruses.^{39–41} Consistent with this observation, it has been shown that the first 70 amino acid residues of HAstV-1 are dispensable for virus-like particle formation using a recombinant baculovirus expression system.²⁷ Jonassen *et al.* also related this arginine-rich region to gene expression regulation processes in other viruses such as coronaviruses, papillomavirus and baculovirus.⁵⁹

Detailed ultrastructure studies of trypsin-treated HAstV-2 virions revealed icosahedral particles with an array of spikes protruding from the surface of the virion.⁴² A low-resolution cryoelectron microscopy image for trypsin-treated HAstV-1 particles showed a solid icosahedral capsid shell 330 Å in diameter and decorated with 30 dimeric spikes extending 50 Å from the surface.⁴

Using the three-dimensional position-specific scoring matrix (3D-PSSM)⁴³ to identify recognizable protein-folding motifs within the astrovirus capsid proteins, meaningful structural predictions are currently providing insights into different functional aspects of the viral

capsid biology. This approach has recently been used by two different research groups, leading to similar conclusions.^{27,35} Consistently, all significant matches for VP34 sequence from different HAstV serotypes and other animal astroviruses correspond to coat proteins from simple, icosahedrally symmetric viruses with jelly-roll β -barrel subunits, such as Theiler's murine encephalomyelitis virus, bean pod mottle virus, carnation mottle virus, and tomato bushy stunt virus. Structural alignments of predicted structural homology comparisons between a fragment of VP34 sequence and coat proteins from these viruses are shown in Fig. 3. Interestingly, the putative RNA-binding domain present at the N-terminus of VP34 was never included in positive alignments, confirming a different biological function for these two domains of the VP34 protein.

The evaluation of protein-folding prediction for the variable domain (amino acid 416 to the end of ORF2) showed significant structural homology to virus-related proteins and receptors, as well as to some non-viral receptor-ligand interactions.³⁵ These results support the idea that this capsid protein domain is responsible for the receptor-interaction process, which dictates cell tropism and may vary enormously between different astrovirus strains.³⁵ Of interest, most significant alignments did not include the C-terminal region of ORF2, suggesting that this sequence would not be included in the infectious particle, as some of the proposed models of morphogenesis suggest.⁸

These prediction results would indicate that capsid proteins from all astroviruses fold in a manner similar to that of well-studied ssRNA icosahedral viruses with $T = 3$ symmetry, and that the predicted β -barrel domain found in VP34 protein could be the building block for capsid assembly. One common feature of ssRNA viruses that show $T = 3$ symmetry is that their capsids are made of a single structural building block. Assuming that 180 copies of the VP34 protein are required to give a particle of 33 nm with $T = 3$ symmetry and that the spikes are formed by VP26 and VP29 proteins would be consistent with the structural prediction that suggests that VP34 is a structural homolog to known icosahedral coat proteins and that the variable domain (VP26/VP29) is involved in receptor binding.

580 Structure-based Study of Viral Replication

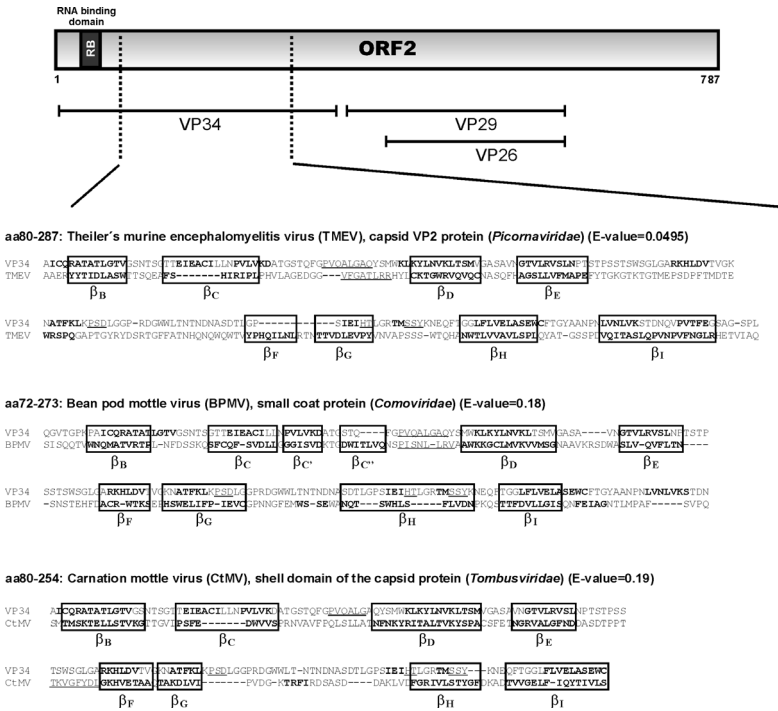


Fig. 3. Schematic representation of the predicted eight β -barrel domain structure within the VP34 capsid protein of HAstV and three structure-based sequence alignments of HAstV sequence with capsid proteins from other viruses: VP2 coat protein of Theiler's murine encephalomyelitis virus (Protein Data Bank 1TMF), small subunit of Bean pod mottle virus coat protein (Protein Data Bank 1PGL), and coat protein (shell domain) of Carnation mottle virus (Protein Data Bank 1OPO). E-value is the measure of confidence in prediction of the 3D-PSSM alignment. Beta-barrel domains are indicated by boxes, beta-strands are indicated in bold, and underlined sequences correspond to alpha-helices. RB denotes the putative RNA binding domain. Amino acid (aa) positions refer to HAstV-1 Oxford reference strain (accession no. L23513).

HAstV Virus-like Particles (VLPs)

Recently, it has been demonstrated that expression of the complete ORF2 of HAstV serotypes 1 and 2 leads to the formation of VLPs using the baculovirus and vaccinia virus recombinant expression systems, respectively.^{25,27} Using the baculovirus system, VLPs can be

obtained either with the expression of a complete ORF2, a truncated form of ORF2 lacking the first 70 amino acids, or after replacing these 70 amino acids with the green fluorescent protein (GFP). Recombinant HAstV VLPs will provide a virtually unlimited supply of highly purified viral capsids with many immediate applications, such as the characterization of antigenic properties of the mature capsid proteins and epitope mapping studies. Additionally, VLPs may also be used to determine the HAstV atomic structure using X-ray crystallographic techniques. Another potential application of the expressed recombinant VLPs is to study cell-virus interactions in binding assays, which may allow the identification of the cellular receptor involved in viral infection.

Interestingly, during the formation of HAstV VLPs in insect cells, a smaller second type of structure consisting of 16 nm ring-like units was observed, mostly after disassembling the 38 nm VLPs through the addition of EDTA.²⁷ As previously described for Norwalk virus as well as for other $T = 3$ RNA plant viruses, these 16 nm VLPs could correspond to structures that are formed when 60 units of the capsid protein assemble into a structure with $T = 1$ symmetry. Depending on factors such as trypsin digestion, the presence and size of the RNA genome, the presence of divalent cations, the ionic strength and the pH, capsid proteins of these viruses are able to self-assemble *in vitro* into $T = 3$ or smaller $T = 1$ structures.^{44,45}

Nonstructural Proteins

The nonstructural proteins of the virus are translated from the genomic viral RNA as two polyproteins, one of which contains only ORF1a (nsP1a: 101 kDa) and the other that includes ORF1a/1b (nsP1a1b: 160 kDa) and is translated via a -1 ribosomal frame-shifting event between ORF1a and ORF1b.¹¹ Both proteins are proteolytically processed, giving rise to a variety of proteins. While nsP1b (59 kDa) corresponds to the viral RNA-dependent RNA polymerase — which can be aligned with other polymerases of the Koonin's supergroup I together with picornaviruses, caliciviruses and certain plant viruses — little is known about the role of most of the mature nsP1a products (Fig. 4). The protease motif has features consistent with chymotrypsin-like

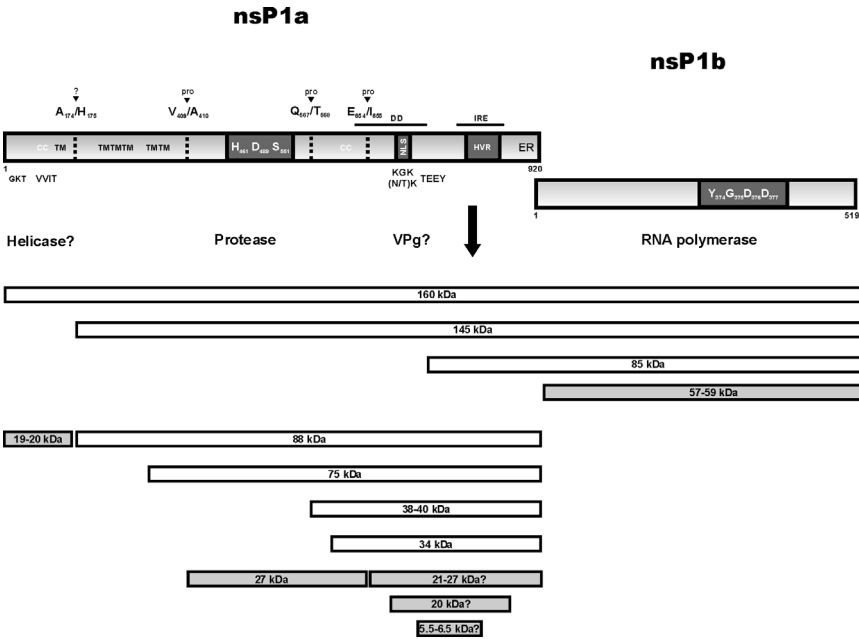


Fig. 4. Schematic depiction of HAstV nonstructural polyproteins nsP1a and nsP1b, and proposed proteolytic processing pathway. Predicted transmembrane helices (TM), nuclear localization signal (NLS), immunoreactive epitope (IRE), coiled-coil structures (CC), predicted death domain (DD), hyper-variable region (HVR), KKXX-like endoplasmic reticulum retention signal (ER), protease motif, RNA polymerase motif, and putative RNA helicase and VPg motifs are shown. Inverted arrowheads denote putative cleavage sites dependent on unknown cellular proteases (?), and the viral protease (pro). White boxes represent intermediate products, while grey boxes represent the final products. See text for details. Amino acid positions are numbered according to HAstV-1 Oxford reference strain (accession no. L23513).

proteases of other plus-stranded RNA viruses, including rabbit hemorrhagic disease virus and feline calicivirus,¹¹ with the substitution of a serine for a cysteine at the third catalytic amino acid residue (His₄₆₁, Asp₄₈₉, and Ser₅₅₁). Some of the amino acid residues involved in substrate binding involve Thr₅₄₆ and His₅₆₆.^{11,46}

Several transmembrane helices (TM) have been identified at the N-terminus of nsP1a. Although the exact number varies among different studies,^{11,20} it is believed that these TMs are responsible for

the anchorage of the RNA replication complex to intracellular membranes. Based on the software predictions on the orientation of the TM helices, it seems likely that the N-terminal region of the helices in nsP1a is on the cytosolic side of a membrane. Sequence analysis has also predicted a KKXX-like endoplasmic retention (ER) signal that suggests an association between the viral RNA replication process and ER-derived intracellular membranes.⁴⁷ In addition, computer prediction led to the identification of a bipartite nuclear localization signal (NLS).¹¹ Although some nsP1a-derived products have been shown to accumulate in the nuclei of infected cells,⁴⁸ the significance of a nuclear involvement in an RNA virus lifecycle is yet to be understood. Recently, two common motifs for astrovirus and pestivirus RNA helicase (GKT and VVIT) have been proposed upstream from the protease coding region. These motifs would represent motifs I and II of the seven conserved motifs of SF2 helicases and would be functionally important in ATP binding.¹⁷

The presence of an immunoreactive epitope (IRE) close to the C-terminus of nsP1a was partially characterized some years ago, indicating the high immunogenicity of this region.⁴⁹ In addition, a hyper-variable region (HVR) has been identified close to this epitope by different authors following genetic characterization of HAstV-1, HAstV-3, HAstV-4, and HAstV-8.^{9,13,47,50} Variability within this HVR has been associated with different viral RNA replication and growth properties, as well as with different virus RNA levels in feces from children with gastroenteritis, suggesting a relationship between certain genotypes and some viral properties related to its pathogenic phenotype.⁵¹ Variability within this region consists mainly of high rates of nucleotide and amino acid substitutions, as well as many insertions and deletions that retain the reading frame. Interestingly, a 15-amino acid deletion identified some years ago was related to adaptation of HAstV to certain cell lines,⁵⁰ and recently this region has been associated with a distinctive RNA replication pattern. Although the molecular mechanisms that regulate the efficient minus and plus RNA strands, including the synthesis of a subgenomic RNA, remain unclear, mutagenesis studies have demonstrated that genetic variability of the C-terminal of nsP1a affects the virus RNA replication phenotype.⁵¹ Additionally, using antibodies against the HVR, it has been shown

that the C-terminal protein of nsP1a colocalizes with the endoplasmic reticulum and viral RNA in HAstV-4 CaCo-2 infected cells, suggesting the involvement of this protein in the RNA replication process in endoplasmic reticulum-derived intracellular membranes.⁴⁷

Computer analysis of nsP1a has recently revealed the presence of two coiled-coil regions (CC) common to all known human astroviruses, which are hypothesized to be involved in the formation of protein oligomers.²⁰ The absence of a methyltransferase-encoding region and the similarity of the HAstV RNA-dependent RNA polymerase motif with primarily polymerases of VPg-containing viruses raise the possibility that the HAstV genome, as well as the subgenomic RNA, may be linked to a VPg protein in its 5'-end. A convincing VPg domain has not been identified, but it may be located upstream from the protease motif, where it has been postulated that Ser₄₂₀ could link the VPg to viral RNA.¹¹ However, based on sequence comparison between human and animal astroviruses with some members of the *Caliciviridae*, *Picornaviridae* and *Potyviridae* families, other authors have also suggested that amino acid Tyr₆₉₃ at the conserved TEEY-like motif could also display a VPg function.^{17,20} Indeed, one of the conserved amino acid motifs characteristic of VPg of caliciviruses [KGGK(N/T)K]⁵² can also be identified upstream from this Tyr₆₉₃ residue at a similar distance as it is found in the calicivirus genome.

Finally, it has recently been reported that products derived from nsP1a lead to apoptosis of the host cell, resulting in efficient virus replication and particle release.^{23,28} The analysis of the secondary structure of the whole astrovirus protein sequence revealed the presence of an optimal six α -helix structure of 95 amino acids close to the C-terminus of nsP1a, which displays homology to the death domain superfamily.²⁸ The presence of a putative death domain in nsP1a suggests a direct link between this polyprotein and the apoptotic pathway.

Proteolytic Processing

The proteolytic processing of the HAstV nsP1a polyprotein (920–935 amino acids, depending on serotype) has been only partially characterized, and some of the reported data are still conflicting.^{40,46,48,53,54}

Data obtained with *in vitro* studies with HAstV-1 suggest that it is cleaved at Gln₅₆₇/Thr₅₆₈ in a process that is dependent on the viral 3C-like serine protease, resulting in an N-terminal p64 fragment that includes the protease motif and a C-terminal p38 product that includes the NLS and the IRE.⁴⁶ However, using a similar approach with a different serotype, Gibson *et al.* could not identify any proteolytic activity,⁵³ suggesting that the viral protease may require some cellular factors.

A later work performed on transfected BHK-21 cells suggested a proteolytic cleavage site close to amino acid 170 of nsP1a, which would generate an N-terminal product of approximately 20 kDa, and two cleavages sites around amino acid 410 and 655, which would result in a 27 kDa protein containing the protease motif. The viral protease would be responsible for only the two latter cleavages.⁴⁰

Performing pulse chase experiments on HAstV-8 infected CaCo-2 cells, Méndez *et al.* identified polypeptides of 88 and 75 kDa, and polypeptides of 145 and 85 kDa as intermediate products of nsP1a and nsP1a1b proteolytic processing, respectively.⁵⁴ As final products, they reported two proteins of 19 and 20 kDa (N-terminal product), two proteins of approximately 27 kDa (protease), and a protein of 57 kDa (RNA polymerase). Results from kinetic experiments were consistent with the idea that a cotranslational processing at the N-terminus of nsP1a and nsP1a1b polyproteins occurs during infection in a process that is independent of the viral protease activity. Based on sequence homology with bovine viral diarrhea virus, the authors suggested that cleavage would occur after the conserved sequence GGYA, between amino acid residues 171 and 174 of nsP1a. Using transient expression experiments on BHK-21 cells, the authors could detect an additional protein of 20 kDa that would correspond to the C-terminal product of nsP1a. Willcocks *et al.* detected products of 75, 34, 20, 6.5, and 5.5 kDa in HAstV-1 infected CaCo-2 cell extracts using antibodies against the 30% C-terminal end of nsP1a, as well as a protein of 59 kDa using antibodies to nsP1b.⁴⁸ Using computer predictions, after mapping the VPg domain downstream the protease domain, Al-Mutairy *et al.* suggested potential cleavage sites for the VPg protein.¹⁷ The authors hypothesized that the N-terminal cleavage

site for astrovirus VPg would be located immediately upstream from the KGK(N/T)K motif, and that its C-terminal cleavage site might be found between 92–143 amino acid residues downstream from the N-terminal cleavage site. However, the actual occurrence of a VPg protein within this region, as well as its boundaries, is yet to be experimentally confirmed.

Finally, immunoprecipitation studies on HAstV-4 infected CaCo-2 cells with an antibody against a synthetic peptide corresponding to the HVR sequence close to the C-terminus of nsP1a led to the detection of five proteins in the range of 21–27 kDa, and at least one of them could be post-translationally modified by phosphorylation on its serine and/or threonine residues.⁴⁷ In addition, after finding proteins larger than expected in nuclear extracts from recombinant baculovirus-infected cells expressing amino acids 643–940 of nsP1a, Willcocks *et al.* also hypothesized post-translational modifications.⁴⁸ Since many proteins destined for the nucleus are heavily glycosylated, the authors suggested that nsP1a proteins are modified by glycosylation.

Structural Computational Predictions

Although it is thought that all nonstructural proteins would form the RNA replication complex and that the transmembrane helices present on nsP1a would anchor the complex to intracellular membranes, the functional properties of most of the individual proteins of HAstV have been poorly characterized. Results obtained after making functional predictions using the 3D-PSSM server for each of the different HAstV uncharacterized non-structural proteins are shown in Table 1. The previously identified sequence with homology to a death domain structure is also included.²⁸ All matches found, except the protease and the polymerase motifs, displayed just a moderate degree of certainty, and the only region that did not display any significant match corresponded to the N-terminal product, which is rich in transmembrane helices. Amino acids 191–353 showed structural homology to the DNA-protecting protein under starved conditions (DPS protein), a ferritin homolog that unspecifically binds and protects DNA⁵⁵; amino acids 596–739 matched a sarcoplasmic calcium-binding

Table 1. Top Protein Fold Matches for HAstV Nonstructural Proteins using 3D-PSSM Web Server

HAstV Region ^a	Result	Superfamily	Template Length (aa)	Sequence Identity (%)	E-value
nsP1a					
191–353	Dodecameric ferritin homolog DPS	DNA protecting protein during starvation (DPS); Ferritin-like	159	10	0.562
370–655	DegP (HtrA)	Heat shock protease HtrA	299	13	2.68e-10
596–739	Sarcoplasmic calcium-binding protein	EF-hand, calmoduline-like	174	9	0.686
607–737	Human Fas	Death domain	127	22	8.52
718–897	Tumor suppressor bin1 (amphiphysin II)	Myc box dependent interacting protein 1; endocytosis/exocytosis	213	18	1.71
nsP1b					
44–515	Rabbit hemorrhagic disease virus RNA-2 dependent RNA polymerase	RNA-directed RNA polymerase	493	13	0.149e-3

^aAmino acid (aa) positions are numbered according to HAstV-1 Oxford reference strain (accession no. L23513).

588 *Structure-based Study of Viral Replication*

protein, a protein with an EF-hand calmoduline motif; and amino acids 718–897 displayed homology to the tumor suppressor bin1 (amphiphysin II), which belongs to a family of adapter proteins that have been implicated in a variety of different cellular processes, including endocytosis, actin cytoskeletal organization, transcription, and stress responses.⁵⁶ Although the sequence corresponding to the HAstV protease showed structural relationship to the protease motif of the heat shock protein DegP (Htra) of *Escherichia coli*, the NCBI conserved domain search identified homology to the Equine arterivirus serine endopeptidase S32 (pfam 05579). Although not contained within the significant alignment, both proteases contain a C-terminal extension domain that may play a role in mediating protein-protein interactions, which may be a novel way of regulating its proteolytic activity.^{57,58}

Using similar approaches, an extensive proteomic computational analysis was performed for HAstV-4 nsP1a C-terminal product.⁴⁷ This analysis indicated that the proportion of residues that were not assigned to a particular secondary structure was higher within this region of nsP1a than in other regions of the genome, revealing the high degree of structural flexibility that this region may have. Accordingly, analysis of genetic variability revealed that this region was the most variable region of the whole nonstructural coding region, showing a high degree of tolerance to insertions and deletions. The authors also suggested a higher occurrence of post-translational modification motifs, such as phosphorylation and O-glycosylation. Although sequence homology searches only displayed proteins with a moderate degree of similarity, the most significant results involved nonstructural proteins from single-stranded RNA viral families involved in the formation of the RNA replicase complexes and/or in the regulation of both viral and cellular transcription processes, and at the superfamily level, the protein was assigned to the “winged-helix” DNA-binding domain superfamily (SSF46785).

HAstV Replication Cycle

Replication of genomes of all well-characterized positive-strand RNA viruses — including plant, animal, and insect viruses — occurs in large

replication complexes associated with intracellular membranes. For HAstV, it has been suggested that RNA replication is associated to membranes derived from the endoplasmic reticulum, at the perinuclear region, and ultrastructural analysis has revealed the presence of viral aggregates close to the nuclear periphery surrounded by a high number of double-membrane vacuoles.⁴⁷ Although the exact composition of these replication complexes is not known, it is likely that all nonstructural proteins are required and that the temporary regulation of protein processing of the replicase complex may play a role in the regulation of the minus- and plus-strand RNA synthesis processes. Of interest, variability within the protein encoded at the C-terminal region of ORF1a has been shown to affect the RNA replication pattern, and a certain functional effect of phosphorylation of the protein has been suggested.⁵¹ Although some nonstructural proteins have been detected by immunofluorescence within the nucleus of infected cells,⁴⁸ the involvement of the nuclear phase in the replication cycle of HAstV is not fully understood.

On the other hand, the coordination of genome packaging into maturing particles has not yet been defined, but the structural determinants for particle assembly seem to be carried out by the structural proteins alone, since it is possible to make VLPs even without the first 70 amino acids of the capsid polyprotein. The fact that structural proteins colocalize with nonstructural proteins suggests the possibility that the membranes might form a scaffold for packaging of the viral RNA into assembling capsids.

Finally, transient expression experiments showed a direct link between HAstV ORF1a encoded proteins and apoptosis induction.²⁸ It seems that astrovirus replication is tightly linked to an apoptotic host cell response, since cellular caspases seem to be responsible for the first step of capsid maturation, before the trypsin cleavages that occur extracellularly.²³ While a premature apoptotic response usually impairs viral infectivity, when enough virus progeny has been generated, apoptosis may facilitate the release of virus to bystander cells, evading the inflammatory response. Kinetic analysis of viral replication, indicated by the synthesis of structural proteins from the subgenomic RNA, and onset of apoptosis showed that host cell

apoptosis would be triggered subsequent to viral replication,²⁸ suggesting that the apoptotic process is a late event in the replication cycle of HAstV.

Future Challenges

Although HAstV infectious particles were discovered almost 30 years ago, limited information is still available on the molecular characteristics of the virus, and its genome organization has not been completely clarified. However, the availability of molecular techniques to study epidemiological features of viral infection has resulted in an extensive database of sequence information, with more than 400 entries in the Genbank, and we are beginning to understand its burden in gastroenteritis worldwide, the extent of its variability, and the main characteristics of immunologic responses to infection. Aside from the use of some cell-adapted HAstV strains that replicate in CaCo-2 cells to high titers, the construction of an infectious full-length clone with serotype 1 has been a critical step in development of genetic systems for study of HAstV.¹⁶ Not only may it be useful for identifying the cleavage sites of both nonstructural polyproteins and the capsid precursor, as well as its biological functions, but it also offers promise as a paradigm for understanding the interactions between viral and host cellular processes. Studies on HAstV capsid structure will provide new insights into the understanding of HAstV genetic and antigenic diversity, as well as the characterization of the immune response, two key steps before the generation of potential vaccines.

Acknowledgments

S. Guix was recipient of an FI fellowship from the Generalitat de Catalunya. We acknowledge the technical expertise of the Serveis Científic-Tècnics of the University of Barcelona. This work was supported in part by grants SP22-CT-2004-502571 from the European Union, 2001/SGR/00098 from the Generalitat de Catalunya, and

the Centre de Referència de Biotecnologia de Catalunya (CeRBa), Generalitat de Catalunya.

References

1. World Health Report. (2004) WHO, Geneva.
2. Taterka JA, Cuff CF, Rubin DH. (1992) Viral gastrointestinal infections. *Gastroenterol Clin North Am* **21**: 303–330.
3. Glass RI, Noel J, Mitchell D, *et al.* (1996) The changing epidemiology of astrovirus-associated gastroenteritis: a review. *Arch Virol (Suppl)* **12**: 287–300.
4. Matsui SM, Greenberg HB. (2001) Astroviruses. In B.N. Fields, D.M. Knipe, P.M. Howley, *et al.* (eds), *Fields virology*, pp.875–893. Lippincott Williams and Wilkins, Philadelphia, Pennsylvania.
5. Walter JE, Mitchell DK. (2003) Astrovirus infection in children. *Curr Opin Infect Dis* **16**: 247–253.
6. Appleton H, Higgins PG. (1975) Viruses and gastroenteritis in infants. *Lancet* **1**: 1297.
7. Madeley CR, Cosgrove BP. (1975) 28 nm particles in faeces in infantile gastroenteritis. *Lancet* **2**: 451–452.
8. Méndez E, Fernández-Luna T, López S, *et al.* (2002) Proteolytic processing of a serotype 8 human astrovirus ORF2 polyprotein. *J Virol* **76**: 7996–8002.
9. Méndez-Toss M, Romero-Guido P, Munguia ME, *et al.* (2000) Molecular analysis of a serotype 8 human astrovirus genome. *J Gen Virol* **81**: 2891–2897.
10. Taylor MB, Walter JE, Berke T, *et al.* (2001) Characterisation of a South African human astrovirus as type 8 by antigenic and genetic analyses. *J Med Virol* **64**: 256–261.
11. Jiang B, Monroe SS, Koonin EV, *et al.* (1993) RNA sequence of astrovirus: distinctive genomic organization and a putative retrovirus-like ribosomal frameshifting signal that directs the viral replicase synthesis. *Proc Natl Acad Sci USA* **90**: 10539–10543.
12. Lewis TL, Greenberg HB, Herrmann JE, *et al.* (1994) Analysis of astrovirus serotype 1 RNA, identification of the viral RNA-dependent RNA polymerase motif, and expression of a viral structural protein. *J Virol* **68**: 77–83.

592 *Structure-based Study of Viral Replication*

13. Oh D, Schreier E. (2001) Molecular characterization of human astroviruses in Germany. *Arch Virol* **146**: 443–455.
14. Willcocks MM, Brown TD, Madeley CR, Carter MJ. (1994) The complete sequence of a human astrovirus. *J Gen Virol* **75**: 1785–1788.
15. Monroe SS, Stine SE, Gorelkin L, *et al.* (1991) Temporal synthesis of proteins and RNAs during human astrovirus infection of cultured cells. *J Virol* **65**: 641–648.
16. Geigenmüller U, Ginzton N, Matsui SM. (1997) Construction of a genome-length cDNA clone for human astrovirus serotype 1 and synthesis of infectious RNA transcripts. *J Virol* **71**: 1713–1717.
17. Al-Mutairy B, Walter JE, Pothen A, Mitchell DK. (2005) Genome prediction of putative genome-linked viral protein (VPg) of astroviruses. *Virus Genes* **31**: 21–30.
18. Sosnovtsev S, Green KY. (1995) RNA transcripts derived from cloned full-length copy of the feline calicivirus genome do not require VPg for infectivity. *Virology* **210**: 383–390.
19. Matsui SM, Kiang D, Ginzton N, *et al.* (2001) Molecular biology of astroviruses: selected highlights. *Novartis Foundation Symposium* **238**: 219–233.
20. Jonassen CM, Jonassen TO, Sveen TM, Grinde B. (2003) Complete genomic sequences of astroviruses from sheep and turkey: comparison with related viruses. *Virus Res* **91**: 195–201.
21. Walter JE, Briggs J, Guerrero ML, *et al.* (2001) Molecular characterization of a novel recombinant strain of human astrovirus associated with gastroenteritis in children. *Arch Virol* **146**: 2357–2367.
22. Bass DM, Qiu S. (2000) Proteolytic processing of the astrovirus capsid. *J Virol* **74**: 1810–1814.
23. Méndez E, Salas-Ocampo E, Arias CF. (2004) Caspases mediate processing of the capsid precursor and cell release of human astroviruses. *J Virol* **78**: 8601–8608.
24. Sánchez-Fauquier A, Carrascosa AL, Carrascosa JL, *et al.* (1994) Characterization of a human astrovirus serotype 2 structural protein (VP26) that contains an epitope involved in virus neutralization. *Virology* **201**: 312–320.
25. Dalton RM, Pastrana EP, Sánchez-Fauquier A. (2003) Vaccinia virus recombinant expressing an 87-kilodalton polyprotein that is sufficient to form astrovirus-like particles. *J Virol* **77**: 9094–9098.

26. Geigenmüller U, Ginzton N, Matsui SM. (2002) Studies on intracellular processing of the capsid protein of human astrovirus serotype 1 in infected cells. *J Gen Virol* **83**: 1691–1695.
27. Caballero S, Guix S, Ribes E, *et al.* (2004) Structural requirements of astrovirus virus-like particles assembled in insect cells. *J Virol* **78**: 13285–13292.
28. Guix S, Bosch A, Ribes E, *et al.* (2004) Apoptosis in astrovirus-infected CaCo-2 cells. *Virology* **319**: 249–261.
29. Al-Molawi N, Beardmore VA, Carter MJ, *et al.* (2003) Caspase-mediated cleavage of the feline calicivirus capsid protein. *J Gen Virol* **84**: 1237–1244.
30. Best SM, Shelton JF, Pomepuy JM, *et al.* (2003) Caspase cleavage of the non-structural protein NS1 mediates replication of Aleutian mink disease virus. *J Virol* **77**: 5305–5312.
31. Eleouet J-F, Slee EA, Saurini F, *et al.* (2000) The viral nucleocapsid protein of transmissible gastroenteritis coronavirus (TGEV) is cleaved by caspase-6 and -7 during TGEV-induced apoptosis. *J Virol* **74**: 3975–3983.
32. Goh PY, Tan YJ, Lim SP, *et al.* (2001) The hepatitis C virus core protein interacts with NS5A and activates its caspase-mediated proteolytic cleavage. *Virology* **290**: 224–236.
33. Grand R, Schmeiser K, Gordon E, *et al.* (2002) Caspase-mediated cleavage of adenovirus early region 1A proteins. *Virology* **301**: 255–271.
34. Zhirnov OP, Konakova TE, Garten W, Klenk HD. (1999) Caspase-dependent N-terminal cleavage of influenza virus nucleocapsid protein in infected cells. *J Virol* **73**: 10158–10163.
35. Krishna NK. (2005) Identification of structural domains involved in astrovirus capsid biology. *Viral Immunol* **18**: 17–26.
36. Wang Q, Kakizawa J, Wenng L, *et al.* (2001) Genetic analysis of the capsid region of astroviruses. *J Med Virol* **64**: 245–255.
37. Bass DM, Upadhyayula U. (1997) Characterization of human serotype 1 astrovirus-neutralizing epitopes. *J Virol* **71**: 8666–8671.
38. Herrmann JE, Hudson RW, Perron-Henry DM, *et al.* (1988) Antigenic characterization of cell-cultivated astrovirus serotypes and development of astrovirus-specific monoclonal antibodies. *J Infect Dis* **158**: 182–185.

594 *Structure-based Study of Viral Replication*

39. Baer ML, Houser F, Loesch-Fries LS, Gehrke L. (1994) Specific RNA binding by amino-terminal peptides of alfalfa mosaic virus coat protein. *EMBO J* **13**: 727–735.
40. Geigenmüller U, Chew T, Ginzton N, Matsui SM. (2002) Processing of nonstructural protein 1a of human astrovirus. *J Virol* **76**: 2003–2008.
41. Geigenmüller-Gnirke U, Nitschko H, Schlesinger S. (1993) Deletion analysis of the capsid protein of Sindbis virus: identification of the RNA binding region. *J Virol* **67**: 1620–1626.
42. Risco C, Carrascosa JL, Pedregosa AM, *et al.* (1995) Ultrastructure of human astrovirus serotype 2. *J Gen Virol* **76**: 2075–2080.
43. Kelley LA, MacCallum RM, Sternberg MJE. (2000) Enhanced Genome Annotation using Structural Profiles in the Program 3D-PSSM. *J Mol Biol* **299**: 499–520.
44. Harrison SC. (2001) Principles of virus structure. In B.N. Fields, D.M. Knipe, P.M. Howley, *et al* (eds), *Fields virology*, pp. 53–85. Lippincott-Raven Publishers, Philadelphia, Pennsylvania.
45. White LJ, Hardy ME, Estes MK. (1997) Biochemical characterization of a smaller form of recombinant Norwalk virus capsids assembled in insect cells. *J Virol* **71**: 8066–8072.
46. Kiang D, Matsui SM. (2002) Proteolytic processing of a human astrovirus nonstructural protein. *J Gen Virol* **83**: 25–34.
47. Guix S, Caballero S, Bosch A, Pintó RM. (2004) C-terminal nsP1a protein of Human Astrovirus colocalizes with the endoplasmic reticulum and viral RNA. *J Virol* **78**: 13627–13636.
48. Willcocks MM, Boxall AS, Carter MJ. (1999) Processing and intracellular location of human astrovirus non-structural proteins. *J Gen Virol* **80**: 2607–2611.
49. Matsui SM, Kim JP, Greenberg HB, *et al.* (1993) Cloning and characterization of human astrovirus immunoreactive epitopes. *J Virol* **67**: 1712–1715.
50. Willcocks MM, Ashton N, Kurtz JB, *et al.* (1994) Cell culture adaptation of astrovirus involves a deletion. *J Virol* **68**: 6057–6058.
51. Guix S, Caballero S, Bosch A, Pintó RM. (2005) Human astrovirus C-terminal nsP1a protein is involved in RNA replication. *Virology* **333**: 124–131.
52. Dunham DM, Jiang X, Berke T, *et al.* (1998) Genomic mapping of a calicivirus VPg. *Arch Virol* **143**: 2421–2430.

53. Gibson CA, Chen J, Monroe SA, Denison MR. (1998) Expression and processing of nonstructural proteins of the human astroviruses. *Adv Exp Med Biol* **440**: 387–391.
54. Méndez E, Salas-Ocampo MP, Munguía ME, Arias CF. (2003) Protein products of the open reading frames encoding nonstructural proteins of human astrovirus serotype 8. *J Virol* **77**: 11378–11384.
55. Grove A, Wilkinson SP. (2005) Differential DNA binding and protection by dimeric and dodecameric forms of the ferritin homolog Dps from *Deinococcus radiodurans*. *J Mol Biol* **347**: 495–508.
56. Muller AJ, Baker JF, DuHadaway JB, *et al.* (2003) Targeted disruption of the murine Bin1/Amphiphysin II gene does not disable endocytosis but results in embryonic cardiomyopathy with aberrant myofibril formation. *Mol Cell Biol* **23**: 4295–4306.
57. Barrette-Ng IH, Ng KK, Mark BL, *et al.* (2002) Structure of arterivirus nsp4. The smallest chymotrypsin-like proteinase with an alpha/beta C-terminal extension and alternate conformations of the oxyanion hole. *J Biol Chem* **18**: 39960–39966.
58. Jeffery CJ. (2004) Molecular mechanisms for multitasking: recent crystal structures of moonlighting proteins. *Curr Opin Struct Biol* **14**: 663–668.
59. Jonassen CM, Jonassen TO, Saif YM, *et al.* (2001) Comparison of capsid sequences from human and animal astroviruses. *J Gen Virol* **82**: 1061–1067.

