Asymmetry in icosahedral viruses
Daniel J Goetschius¹, Colin R Parrish² and Susan Hafenstein¹,³

Although icosahedral viruses have obvious and highly symmetrical features, asymmetric structural elements are also present. Asymmetric features may be inherent since the genome and location of minor capsid proteins are typically incorporated without adhering to icosahedral symmetry. Asymmetry also develops during the virus life cycle in order to accomplish key functions such as genome packaging, release, and organization. However, resolving asymmetric features complicates image processing during single-particle cryoEM analysis. This review summarizes the current state of knowledge regarding asymmetric structural features with specific examples drawn from members of picornaviridae, paroviridae, microviridae, and leviridae.

Addresses
1 Department of Biochemistry and Molecular Biology, Penn State University, W231 Millennium Science Complex, University Park, PA 16802, USA
2 Baker Institute for Animal Health, Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
3 Department of Medicine, Penn State University College of Medicine, Hershey, PA 17033 USA

Corresponding author: Hafenstein, Susan (shafenstein@psu.edu)

This review comes from a themed issue on Virus structure and expression
Edited by Juliana Reis Cortines and Peter Prevelige Jr
For a complete overview see the Issue and the Editorial
Available online 28th June 2019
https://doi.org/10.1016/j.coviro.2019.05.006
1879-6257 © 2019 Elsevier B.V. All rights reserved.

Summary
A feature of many viruses is their apparently symmetrical shape, which includes the icosahedral capsids of many non-enveloped viruses, the capsids packaged within some enveloped viruses, and the glycoproteins displayed on the outside of other enveloped viruses. This apparent symmetry has typically been emphasized during the structural analysis, since the icosahedral averaging imposed resulted in the loss of any unique features, which in crystallographic studies prevented crystallization. Most bacteriophages have large and obvious unique tails, receptor-binding proteins, or portals that were accounted for from the earliest studies by separately solving the structures to deal with the symmetry mismatches. However, in the smaller icosahedral viruses, portal structures, packaged DNA or RNA genomes, low copy number proteins, or other types of asymmetry have been nearly impossible to study using a structural approach and generally remain poorly understood.

Recent technological advances, particularly in cryogenic electron microscopy (cryoEM), have identified asymmetric features previously missed. It now appears likely that all viruses contain asymmetric and submolar structures that play important functions in many aspects of the virus replication cycles and that many of these will be key targets for the development of new antiviral drugs. Here, we review the background to these structures in a number of animal viruses and review some of the updated functions that are being revealed. Given the rapid development of new technologies there is no doubt that we are only just starting to reveal these functions, and we can expect to see many new and often un-anticipated structures and functions in the near future.

Structures of icosahedral viral capsids and symmetry averaging
From the earliest use of electron microscopy to study viruses it was apparent that many viral capsids appeared spherical, and that once the genomic information was obtained it was clear that the viral capsids must be assembled from multiple copies of a smaller number of viral proteins that made up the capsid subunits. Early work by Caspar and Klug established the principle of ‘quasi-equivalence’, which is the foundation for modern structural virology [1]. The most common form for spherical viruses, the icosahedron has fivefold, threefold and twofold rotational symmetry axes and the icosahedral lattice presents the most efficient organization of repeating subunits. The simplest architecture comprises 12 pentamers and larger icosahedral shells are constructed with additional hexamers in arrangements that depend on quasi-equivalent interactions due to variable environments for the chemically identical structural proteins. Virus icosahedral symmetry is defined by the triangulation number, or T number, which describes how many subunits make up the capsid. Assembling the capsid from many of the same building blocks, allows the virus to package a relatively small genome that encodes the main structural proteins, which are expressed in multiple copies. The structural biologist has also benefited by using this structural redundancy to gain resolution by averaging all the many subunits together to increase the available data.
Virus capsids range from the most simple icosahedral symmetry of 60 repeating subunits ($T = 1$) to giant viruses such as Cafeteria roenbergensis that has a $T$ number of 499 [2]. Since there is a direct relationship between the particle number and resolution of the map, just acquiring 100 icosahedral virus capsids can provide important information about the overall virus structure. However, averaging together all of the capsid subunits assumes that each subunit is identical to the others. If there is any unique aspect to the virus capsid, that information would be lost during the reconstruction process.

Since the inception in 1974 [3], cryo EM has made steady progress; however, the recent transformative advances in sample preparation, data collection, and data processing have brought us into a new era of atomic resolution. In the field of cryoEM image reconstruction, icosahedral viruses have played a leading role due to the symmetry displayed by the repeating capsid protein subunits. However, this traditional capsid-wide averaging approach has obscured unique features of icosahedral viruses throughout the history of cryoEM, as can be illustrated well by the example of the herpesvirus capsid (Figure 1).

**Asymmetry in icosahedral viruses**

Every icosahedral virus capsid has inherent asymmetry upon incorporating the genome. Whether or not each of the repeating subunits that makes up the capsid shell are identical, the genome with or without its packaging proteins (e.g. nucleoproteins) does not completely follow the symmetry of the capsid, but folds into an energetically favorable conformation allowed by interactions with the internal capsid surface. The viral nucleoproteins, N-terminal extensions of capsid structural proteins, or charged polymers such as polyamines may interact specifically or non-specifically with the packaged genome to neutralize the negative charge and aid in the packaging process. The result is that parts of the genome may become partially icosahedrally ordered and thus visualized during the reconstruction process. The X-ray crystal structure of Panicum Mosaic Virus (PMV) revealed 17 nt hairpins of genomic RNA associated to the interior of the capsid at each of the 60 equivalent subunit interfaces sites. In satellite tobacco mosaic virus (STMV), $80\%$ of the 1058 nucleotides forms stem-loop elements at each icosahedral edge [4]. Similar stem-loop arrangements have been proposed for southern bean mosaic virus [5] and observed in the structure of bean pod mottle virus, which orders $20\%$ of the genome [6]. Ordering of the genome is also present for the ssDNA-containing parvovirus, where icosahedrally ordered portions of the canine parvovirus (CPV) or minute virus of mice (MVM) genomes were seen to form unusual loop conformations, with the bases pointing outwards toward the interior of the capsid, and the phosphates surrounding metal ions on the inside. The coat protein interacts with the DNA bases at the

**Figure 1**

Incremental cryoEM advances improved the resolution so that the hand of the herpes virus capsid was assigned correctly in 2002, but it was another 4 years afterward that the inherent asymmetry was discovered. CryoEM through the years as illustrated by herpesvirus: (a–l) Resolution increasing as cryoEM technique and data collection approaches improve from 1990–2000 resolution improved from 45Å to 8.5Å. (a) Baker 1990, 45Å [46], (b) Newcomb 1993, resolution not provided [47], (c) Conway 1996, 24Å [48], (d) Zhou 1999, 20Å [49], (e) Zhou 2000 8.5Å [50], (f) With Cheng 2002 at 18Å, the correct handedness was reported for the first time [51], (g) Baker 2006, 8.5Å [52], (h & i) Asymmetric approaches visualized a unique portal for packaging of DNA. This portal had been obscured until 2006 due to symmetry averaging. Portal visualized by Cardone (not shown) and Chang 2007 [53,54], and Rochat [55] (j) Schmid 2012 visualized the portal with a tail-like structure using an asymmetric approach [56], (k & l) Using advances in hardware and software approaches the current locally refined map of herpesvirus reached an atomic resolution of 3.5Å [57], with an asymmetric map at 8Å featuring the unique portal [58]. Thus early approaches that had to rely on symmetry averaging, effectively “hid” from discovery the asymmetric aspect of herpesvirus: a unique portal.
icosahedral twofold [7], suggesting the existence of a genomic DNA-recognition site within the parvovirus capsids [8].

There are two basic models for packaging the RNA or DNA genome. In the first an empty capsid is first formed, called the procapsid, into which a progeny genome is packaged. This process uses energy, specific translocation machinery, and requires the use of a single packaging portal (which may be structurally distinct) and the formation of an asymmetric packaging complex. In the second model, virus assembly subunits are recruited to and assemble around the progeny genome to form the capsid and incorporate the genome. This type of assembly most likely occurs at the site of genome replication since as the progeny genome is synthesized, structural elements form to direct the assembly of the virus capsid (Figure 2).

Nucleic acid packaging and portals
Packaging may occur concurrently with genome replication as has been shown for the Microviridae and Parvoviridae. During Microvirus DNA packaging, the packaging/ssDNA replication complex, consisting of replicative form DNA, host cell and viral proteins [9–11], docks to the viral procapsid. Although this complex has yet to be structurally visualized, the results of genetic analyses indicate that docking occurs within a groove in the capsid that spans the twofold axis of symmetry [12]. The DNA likely enters the procapsid through one of the pores at the

Figure 2

Asymmetric features of viruses by genus. Genome packaging, host receptor recognition, and genome release are fundamentally asymmetric processes within the lifecycle of the virus, often reflected in asymmetric features on the mature virus. (A) In picornaviridae, G<sub>ATPase</sub> mediates genome packaging [50,51]. The mature virus contains a combination of VP0 and the genome-associated VPg [52,53]. Host recognition is an asymmetric event, represented here by CVB3 and the coxsackie-adenovirus receptor [29]. Genome release likely occurs at a twofold pore [28]. (B) In parvoviridae, NS1 both initiates DNA synthesis and helps to package the genome [54]. VP2 represents the majority of capsid proteins, substituted asymmetrically by fewer than 12 VP1 molecules [22,23]. Host recognition is fundamentally asymmetric, represented here by canine parvovirus and the transferrin receptor [56,64]. Genome release is through a unique fivefold pore [44]. (C) In microviridae such as X174, viral proteins A and C interact with the host rep helicase to replicate and package the viral genome at threefold symmetry axis [65]. Symmetry-breaking features of the mature virus include ~10 asymmetrically incorporated copies of structural protein H [66]. In X174, host-recognition and genome release occur as a dramatic remodeling event at the fivefold vertex, with the formation of a portal piercing the bacterial cell wall [55,41,42]. (D) in leviviridae (MS2, Q<sub>t</sub>), genome encapsidation occurs when the capsid forms around the genome [67]. A single capsid protein is substituted by maturation protein A [17–19], which functions in recognition of the bacterial F-pilus and genome release [16,45].

threefold axis of symmetry, the only location wide enough to accommodate the passage of the two antiparallel, ssDNA strands of the circular genome [13].

Parvoviral packaging of the linear ssDNA into the preformed capsid requires the activities of the large non-structural protein, NS1, which attaches to the 5′-end of the DNA, and functions as a replication initiator, as well as being a 3′–5′ helicase which translocates DNA into empty parvovirus particles. The site of genome packaging is one of the fivefold vertices, and newly produced viruses have ~24 nucleotides of the packaged DNA genome remaining exposed to the exterior, still linked to the NS1 protein via a Tyr residue [14].

Assembly around the viral genome

The highly ordered genome of STMV probably represents the far end of the spectrum for building a capsid around the progeny RNA/DNA. Assembly is driven by dimers of the coat protein interacting with stem loops of progeny RNA, and that is predicted to grow into an aggregate that then rearranges into an icosahedron by fluidly shifting and forming/ releasing interactions [4]. Recent findings suggest that the secondary structure of the RNA differs significantly when comparing the packaged and unpackaged forms [15]. In another example, the small genome of the leiviruses (3400–4300 nucleotides) also mediates coalescence of the capsid proteins around the genome in order to build the virion. A specific RNA hairpin called the ‘operator’ selectively binds to dimers of the coat protein (CP) to drive selection of genome. The first asymmetric reconstruction of the MS2 capsid by tomography showed that the genomic RNA forms a specific folded structure once fully incorporated into the capsid and that the maturation (or A protein) replaces a single coat protein dimer [16]. Subsequent asymmetric studies using high resolution cryoEM have also verified this capsid asymmetry in the related Q-Beta and MS2 phage [17–19].

Locations of minor capsid proteins

Besides the asymmetry of genome packaging, there are ‘minor’ structural proteins, which are present in smaller numbers per capsid than the proteins that assemble the main capsid shell. Minor structural proteins are often incorporated asymmetrically in the mature virus. During assembly the microvirus phiX174 incorporates 10–12 copies of H protein; however, the exact location of the H protein molecules in the mature capsid is unclear and likely the proteins are disordered or asymmetrically incorporated. In the picornaviruses, a single copy of VPg (viral protein genome-linked) is covalently linked to the five prime ends of the progeny genome, where it can act as a primer for transcription (replication) by RNA polymerase upon infection. Although the parvovirus capsid is made up of ~90% VP2, around 6–8 copies of a longer form of the VP2 protein, called VP1, are also present [20,21]. The presence of VP1 is essential to infectivity, and the N terminus of VP1 appears to be sequestered inside the capsids of most parvoviruses, but is extruded from the capsid during entry into the cell. None of these minor capsid proteins have yet been well resolved in the mature capsid using a structural approach.

Asymmetry that develops during the virus life cycle

After assembly and genome packaging, some viruses undergo further maturation processes that induce asymmetry to the capsid. In parvoviruses, VP2 undergoes asymmetric cleavage events that are essential to key steps in entry and infection, perhaps involving host receptor recognition and genome uncoating [22,23]. Picornavirus undergo a metastable phenomenon called ‘breathing’ during which reversible conformational changes take place to externalize sections of VP1 and VP4. Structural studies using antibodies to capture VP1 showed that less than 60 copies of these normally internal epitopes are exposed, establishing that breathing of the icosahedral capsid is asymmetric [24,25].

In a physiological setting, viruses recognize distinct molecules (receptors) present on a permissive host cell for attachment and entry, and the initial contact initiates with binding of only one or a few receptors to a focused, local site on the capsid. That interaction is specific and involves key residues such that a single point mutation on either the virus or receptor can alter or completely abrogate binding. Canine parvovirus recognizes transferrin receptor type-1 (TfR) in a highly specific interaction between the apical domain of TfR and the threefold spike of the virus capsid. The interaction is asymmetric when capsid attaches and enters cells; furthermore, fewer than four TfR molecules can bind each virus in situ, even when excess receptor is presented in solution.

For host entry most picornaviruses bind receptors at a cleft on the capsid surface called the canyon. Currently there are two chief models for the selection of a single site on the virus capsid that will allow subsequent successful release of the genome. In one, a unique site on the capsid is established by the recognition of the receptor, which triggers the formation of an asymmetric entry intermediate, also known as the altered particle or A-particle. In this model, it is the initial recognition site and asymmetric interaction with the host that drive the formation of the unique portal. Alternatively, inherent asymmetry of the picornavirus capsid itself dictates the asymmetric site of the genome-release portal, and receptor recognition simply drives the conformational change at the predeter- mined location. Both models require the input of energy to allow the structural changes required for portal formation and it has been shown that temperature for portal formation is a critical component with receptor recognition serving as a key trigger [26].
Several structural studies have sought to mimic the asymmetric interaction with the host in a physiologically relevant manner by using receptor-decorated liposomes or nanodiscs to form asymmetric entry intermediates [27–29]. In vitro, the most common way to trigger picornavirus to form A-particles is by heating to 50°C or by exposing the virus to low pH; however, these approaches apply a global (capsid wide) stimulus to the particle. There have been many moderate and high resolution structures solved using such capsid wide in vitro approaches to create A-particles, although all have been solved by imposing icosahedral symmetry [26,30–34]. Icosahedral refinement of viruses with asymmetric features can attain atomic resolution albeit, by averaging and obliterating the density of the unique structure. Thus, structural studies so far have not differentiated between the host interaction or inherent asymmetric models of portal formation and genome release.

**Genome release**

A single fivefold vertex of phiX174 interacts with LPS initiating the conformational changes required to establish a unique vertex and begin the process to release the genome across the cell wall into the bacterial host [35]. It has been known for decades that phiX174 attaches to the host cell surface where it undergoes a conformational change to form an ‘eclipsed’ particle that subsequently allows DNA release across the host cell wall. In vitro eclipse can be triggered by exposing particles to Lipopolysaccharide (LPS) or metal ions such as Ca²⁺ [36–38]. Calcium eclipsed particles lose genome through the fivefold pore [39,40]. A novel function has been discovered for the asymmetrically incorporated protein H, ten copies of which self-assemble into a tube that is long enough to span the host’s cell wall and wide enough, 22 Å, to accommodate passage of the circular DNA genome [41]. Thus the tail-less microvirus packs a self-assembling conduit that forms an asymmetric appendage to use as a DNA delivery system into the host [42].

The parvovirus genome appears to be at least partially released in a 3′–5′’ direction through a unique fivefold pore, in this case this likely happens in the nucleus, and it is possible that after the 3′-terminal hairpin folds back to form a site for initiation of DNA polymerization [43,44]. It is clear that genome delivery and release rely on a rearrangement of some capsid structures, and the selection of a unique pore for release is likely guided by the position of the genome 3′ terminus. Divalent cations (likely Ca²⁺ or possible Mg²⁺) are incorporated into the capsids, and may be removed by low pH such as would be encountered in the endosome, which may be one trigger for release, perhaps along with receptor binding. The many steps in infection that lead to transport of the genome to the nucleus involve both viral and cellular components.

The MS2 phage has an icosahedral capsid that binds to the bacterial F-pilus with an interaction mediated by the maturation protein that is incorporated into one position in the capsid. A single fivefold vertex interacts with the pilus and undergoes density rearrangements that are attributed to the RNA genome, suggesting that the interaction triggers conformational changes in genome preliminary to release [45]. When bound capsids were analyzed without imposing any symmetry, they were seen to be remarkably asymmetric, showing the position of the unique A, or maturation protein, and the rearrangements of the genome [16].

**Summary**

Although for a long time we have known that there are many important functions controlled by asymmetric structures in icosahedral viruses it has been difficult to analyze at a structural level. The development of key new technologies has made it possible to begin to gain high enough resolution to provide a detailed analysis.

**Conflict of interest statement**

Nothing declared.

**Acknowledgements**

This work was funded in part by the Pennsylvania Department of Health CURE funds, Huck Institutes at The Pennsylvania State University, and by the Office of the Director, National Institutes of Health, under Award Numbers 1 R01 AI107121-01 (SIB), R01 AI092571 (CRP). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**References**


