

Hepatitis A Virus Capsid Structure

David I. Stuart^{1,2}, Jingshan Ren¹, Xiangxi Wang³, Zihao Rao^{3,4} and Elizabeth E. Fry^{1,†}

¹Division of Structural Biology, University of Oxford, The Henry Wellcome Building for Genomic Medicine, Headington, Oxford, UK, ²Diamond Light Sources, Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK, ³National Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing, 100101, China, ⁴Laboratory of Structural Biology, School of Medicine, Tsinghua University, Beijing, 100084, China.

[†] Corresponding Author

XW: e-mail: xiangxiwang@163.com

JR: Tel: +44 1865287548 Fax: +44 1865287546 e-mail: ren@strubi.ox.ac.uk

DIS: Tel: +44 1865287567 Fax: +44 1865 287546 e-mail: dave@strubi.ox.ac.uk

ZR: e-mail: raozh@mail.tsinghua.edu.cn

EEF: Tel: +44 1865287549 Fax: +44 1865287546 e-mail: liz@strubi.ox.ac.uk

Short Title: HAV Capsid Structure

Abstract

Hepatitis A virus (HAV) has been enigmatic, evading detailed structural analysis for many years. Its recently determined high-resolution structure revealed an angular surface without the indentations often characteristic of receptor binding sites. The VP1 β -barrel shows no sign of a pocket factor and the N-terminus of VP2 displays a 'domain-swap' across the pentamer interface, as in a subset of mammalian picornaviruses and insect picorna-like viruses. Structure-based phylogeny confirms this placement. These differences suggest an uncoating mechanism distinct from that of enteroviruses. An empty capsid structure reveals internal differences in VP0 and the VP1 N-terminus connected with particle maturation. An HAV/Fab complex structure, where the Fab appears to act as a receptor-mimic, clarifies some historical epitope mapping data, but some remains difficult to reconcile. We still have little idea of the structural features of enveloped HAV particles.

Introduction

The primary function of the picornavirus outer protein shell (capsid) is to protect the genome and evade immune recognition to enable the transport of the genome into the cellular cytoplasm where replication can be initiated. The structure of the capsid can help us to understand some aspects of this process: how capsid assembly is directed and controlled, how the genome is encapsidated, how interactions with the host cell are coordinated, how the capsid disassembles to facilitate genome replication and how the capsid organization responds to and evades the antiviral response mounted by the host.

The *Picornaviridae* are an extremely well-studied family of viruses, however the *Hepatovirus* genus of which hepatitis A virus (HAV) is the type species are somewhat outliers from, and much less well studied than, for instance, the enteroviruses (e.g. poliovirus and rhinoviruses). Picornaviruses generally possess a non-enveloped 30 nm diameter capsid that packages a single-stranded RNA genome. The genomic RNA contains a single large open reading frame, flanked on each end by relatively short nontranslated RNA segments. The open reading frame encodes a single large polyprotein that is processed by a viral protease into three polypeptide intermediates, P1-P3. P1 is subsequently further processed into 3 structural proteins, VP0, VP3 and VP1, which self-assemble into a spherical capsid with icosahedral (532) symmetry (Fig. 1a). VP0, VP3 and VP1 are roughly similar in shape and arranged so as to approximate T=3 quasi-equivalence (Rossmann and Johnson 1989). In many picornaviruses VP0 undergoes a final maturation cleavage after packaging of the RNA genome, by a poorly understood mechanism, perhaps catalyzed by RNA. This yields the mature VP2 capsid protein and a smaller N-terminal fragment, VP4.

The general structure of the picornaviral capsid has been well characterized. The three largest capsid proteins, VP1-3, each adopt an eight-stranded anti-parallel 'β-barrel' fold in which the N-terminal amino sequences are positioned on the inner surface and the C-terminal residues face externally. The small VP4 protein is located internally and not well visualized in crystallographic models, either because it is flexible or because it contacts the single molecule of genomic RNA which does not possess strict icosahedral symmetry. Five copies of the VP1 capsid protein surround the icosahedral five-fold-axes. Three copies of VP2 and VP3 alternate at the three-fold-axes, and two copies of VP2 abut each other at the two-fold-axes.

Interactions of the external surface of the assembled capsid with a receptor expressed on the surface of the host cell play a critical role in cellular entry of the virus. The surface of the capsid also interacts with antibodies generated by the host's immune system, potentially resulting in neutralization of infectivity. The peptide loops on the surface of the capsid have little structural constraint, and thus tolerate deletions, insertions, or amino acid substitutions that enable the virus to evade antibodies. These surface loops therefore vary significantly in length and structure, so that different picornavirus species may have distinctive shapes. Thus, viruses of one major genus – *Enterovirus* – have a depression (canyon) around the five-fold-vertices, which is highly conserved and frequently harbours the site of receptor interaction (Rossmann et al. 1985). Close to this is an opening into a hydrophobic pocket in the interior of the VP1 β-barrel in which a lipidic molecule (pocket-factor) binds (Smith et al. 1986). In enteroviruses, it seems that the lipid is dislodged from the pocket, leading to its collapse, after binding of the capsid to its receptor and concomitant uptake into early endosomes. These conformational re-arrangements in the capsid result in channels opening up both on and adjacent to the 2-fold-axes that allow VP4 and the N-terminus of VP1, both hydrophobic and positioned internally, to externalize and puncture the endosomal membrane (Ren et al. 2013). This triggers release of the viral RNA, which enters the cytoplasm where it undergoes translation on ribosomes and initiates the process of replication (Wang et al. 2012). Other genera of the *Picornaviridae*

do not possess the VP1 binding pocket, and the site of receptor binding can vary. It is quite possible that the mechanism of uncoating varies with the site of receptor binding, but the details are less well established.

HAV has markedly different physico-chemical properties to nearly all other picornaviruses (Siegl et al. 1981). It is remarkably stable both genetically and physically, being able to withstand temperatures up to 80°C and pH as low as 2, and empty particles appear even more robust than full particles (Wang et al. 2015). This makes disinfection tricky and the process of uncoating, for now, unfathomable. Quite recently it was shown that the virus exists in an enveloped form (garnering membrane from the host to form the so-called eHAV particle) whilst outside cells, potentially shielding itself from immune surveillance (Feng et al. 2013). This perhaps starts to blur the distinction between enveloped and unenveloped viruses and raises the question of whether this is unique to HAV, or occurs also with some other picornaviruses, and has simply escaped attention. HAV also retains the ability to undergo transcytosis across cells when complexed with IgA antibody (Dotzauer et al. 2005; Counihan and Anderson 2016). HAV appears to have a very short (23 amino acids), non-myristoylated VP4 (Shukla et al. 2014). Unlike other picornaviruses, on capsid assembly a 67 residue polypeptide remains attached as a C-terminal extension of VP1. By comparison with other picornaviruses this would have been expected to belong to P2 (protein 2A in the other viruses) thus it is known as VP1-2A or pX (see Chapter 5). pX is likely to mediate the association with membrane envelope, and indeed is later cleaved off by host proteases to yield the mature, nonenveloped capsid (Feng et al. 2013). HAV isolates belong to a single serotype, and this lack of antigenic variation has been attributed to its need to conserve rare codons (Sanchez-Eugenio et al. 2016). These differences made the determination of its structure an imperative as it was expected that the capsid would bear distinguishing features to the structures from other genera and might shed light on its uncoating mechanism. Indeed the analysis of HAV described below provided the first structural information on the *Hepatovirus* genus.

Overall Structure of HAV

The structure was solved for HAV genotype TZ84, cultivated in 2BS cells and inactivated with formaldehyde from the Sinovac company (Wang et al. 2015). Very thin, plate-like crystals were grown for both the mature and natural empty virus (capsids which did not package RNA) and which were separated during sucrose density gradient purification. The crystals were grown in typical small drops from 30% 2-methyl-2,4-pentanediol (MPD) and probably for this reason *in situ* data collection at room temperature (Axford et al. 2012) was impossible since when the X-ray shutter was opened the crystal would twitch out of the beam [see supplementary movie in Wang et al., 2015 (Wang et al. 2015)]. Data were therefore collected from frozen crystals on beamlines I03 and I24 at Diamond Light Source, UK and structures were solved for the mature and empty capsids at 3.5 Å and 3.0 Å respectively. Unlike some enteroviruses, such as EV71 (Wang et al. 2012), the empty particles were essentially indistinguishable in exterior (and therefore antigenic) structure to the mature virus particle. As expected from previous experience with EV71 (Wang et al. 2012) there was no evidence for structural modification attributable to formaldehyde inactivation.

In terms of its broad surface architecture, the HAV particle most closely resembles the smooth, spherical aphthoviruses, although HAV has a noticeably faceted appearance, akin to a triakis icosahedron (Fig 1b). These flat facets mean that there is no canyon. Compared to a typical enterovirus (Rossmann et al. 1985), a shortening of the VP1 BC loop lowers the north wall of the canyon while reductions in the VP2 EF and VP1 GH loops ablate the south wall (Fig. 2a). Compared to the more spherical foot-and-mouth-disease virus (Acharya et al. 1989), the loops at the 5-fold and 3-fold axes in HAV are slightly raised to make the facets. There are no major depressions anywhere on the surface that might immediately suggest themselves as

receptor attachment sites. The HAV receptor story remains incomplete, there was early evidence for the involvement of glycoporphin A (Sanchez et al. 2004) as the erythrocyte receptor since HAV agglutinates red blood cells at acidic pHs. The epitope for glycoporphin interactions was proposed to involve residue Lys 221 of VP1 and the surrounding VP3 residues (102-121) lining the “capsid pit”. This is mapped to the structure in Fig. 2b. Whilst there is no canyon or pit depression, the VP3 residues do not appear to be accessible in the mature capsid. There have also been reports that TIM1 acts as a receptor (Kaplan et al. 1996). However recently, it has been shown to play no role in cellular entry for either quasi-enveloped or naked HAV although it may facilitate early stages of infection by binding phosphatidylserine on the eHAV surface (Das et al. 2017). We believe the recently determined structure of a complex of HAV with a neutralizing antibody may throw some light on receptor binding, and this is discussed below.

In-line with the low buoyant density in CsCl (Siegl et al. 1981) there are no apertures in the capsid to permit the entry of Cs^+ ions. In both the mature and empty capsids, approximately 3 residues are disordered at the C-terminus of VP1 (where pX would have previously have been joined to VP1). No extension is visible but its position on the outer capsid can be inferred. Interestingly, since the structure of HAV was determined, structures of other picornaviruses have been determined that bear an extension at the C-terminus of VP1 which is, more or less well ordered on the surface of the capsid, an example of this is Ljungan virus (LV), where 20 Å high protrusions are seen, displaced some 45 Å from the icosahedral five-fold axes, composed of the 55 C-terminal residues (Zhu et al. 2015)(Fig. 2c).

Unlike enteroviruses (but in common with other picornavirus genera), HAV harbors no contiguous hydrophobic pocket in the VP1 β -barrel. The β -barrel is compressed compared to enteroviruses (Plevka et al. 2013; De Colibus et al. 2014), and the remaining space largely filled with hydrophobic side-chains, as in aphtho and cardioviruses (Acharya et al. 1989) (Fig. 2d). In addition, lengthened βC and βH strands in VP1 essentially cover what would be the entrance to the pocket. It therefore seems unlikely that HAV would be able to bind small molecules within this pocket as observed in enteroviruses and such a line of attack for antiviral compound discovery is unlikely to be fruitful (De Colibus et al. 2014).

Re-wiring of the pentamer boundary structure

One of the most significant differences between HAV and the earlier characterised mammalian picornavirus structures is a change in the structure at the icosahedral two-fold axes. This comprises a swap in position for the N-terminal 52 residues of VP2. The net result is that the overall structural appearance is almost indistinguishable, since the same β -hairpin structure is formed at the two-fold symmetric position, however this hairpin comes from a different chain. This is achieved by a flip in the ψ torsion angle of the peptide of residue 53, which lies very close to the 2-fold axis (only ~5 Å from it). This flip places the first 53 amino acid residues across the inter-pentamer boundary (Fig. 3a), where they form a strong interaction with the adjacent pentamer. As a result of this swap in the N-terminal structures, an extra strand is formed on the VP2 β -barrel of the neighboring pentamer. Similar structural domain swaps have been observed in certain protein structures (Liu and Eisenberg 2002). It results in a very similar overall structure in the case of HAV, but with complex changes in connectivity and in how the adjacent protomers of one pentamer are sewn together with the adjacent pentamer (Fig. 3a). This arrangement was not known to occur in picornaviruses when first recognized in HAV. However, it exists in distantly related picorna-like viruses of insects including Cricket paralysis virus, CrPV (Tate et al. 1999), a member of the *Cripavirus* genus in the family *Dicistroviridae*. Furthermore, structures of several picornaviruses that harbour an uncleaved VP0 have since been solved, including Ljungan virus (LV) (Zhu et al. 2015) and human Parechovirus-1 (HPEV-1)(Kalynych

et al. 2015), and have been found to have exactly the same VP0 swap (Fig. 3a). The domain-swapped N-terminal conformation for VP2 is actually much closer to the N-terminal conformation of VP3 and VP1. This further strengthens the suggestion that HAV belongs to a major clade of picornaviruses that are more 'primitive' and are the most closely related to the insect picorna-like viruses such as CrPV (Tate et al. 1999) and Triatoma virus (TrV) (Squires et al. 2013), also a cripavirus (Fig. 3b).

The mystery of uncoating

The extent of the structural differences detailed above, strongly suggest that HAV uncoats via a mechanism distinct from enteroviruses. However, it also seems to be different from, for instance, the aphthoviruses, where there is a propensity for the capsid to fall apart into pentamers, whereas HAV is almost bullet-proof. Does comparison of the structures of the mature virus and natural empty capsids throw any light on this? The two structures are almost identical externally, but internally the empty capsid has more disordered protein in the vicinity of the three-fold axes (Fig. 3c), corresponding to regions of VP0 (the first 40 residues) and neighboring stretches of the N-terminus of VP1 (the first 47 residues). In the RNA-containing full particles, VP0 is at least partially cleaved and VP4 can be detected by SDS PAGE, although, it is not visible in the structure (Wang et al. 2015). A greater proportion of VP0 appears to be uncleaved in HAV full particles than in other picornaviruses. This is consistent with previous suggestions that the VP0 cleavage is particularly slow (Probst et al. 1999). Empty HAV particles contain VP0, and appear to be similar to empty particles that are often observed in other picornavirus infections, where, in the absence of RNA there is a tendency for the internal structure to be less well ordered. It seems unlikely that such empty particles can encapsidate RNA and they are therefore, probably not on the route to assembly of full particles.

In the full particles, residues 30-39 at the N-terminus of VP1 on the interior of the capsid are disordered. It is notable that this stretch is rich in Lysine residues, which suggests that this region may interact with the RNA (the latter does not adhere to strict icosahedral symmetry and is not usually visible in structures where icosahedral averaging has been used to improve the phases). The use of a positively charged stretch of residues at the N-terminus of a major capsid protein has also been observed in LV and human parechoviruses, although in that case the protein is VP3 rather than VP1 and the interaction with RNA is confirmed by the detailed visualization of sequence specific interactions with packaging sequences (Zhu et al. 2015; Shakeel et al. 2017). Nevertheless, this suggests that this combination of charge interactions in addition to sequence specific interactions may be a characteristic of this clade of picornaviruses. The sedimentation coefficients are *circa* 144S and 82S for the full and natural empty particles respectively (144S is a little less than the 155S expected, for example for a full enterovirus particle) (Wang et al. 2012). We have investigated the robustness of both empty and full particles, and both are extraordinarily tough compared to other picornaviruses (remaining stable at up to 80 °C and at pH's down to about 2) (Walter et al. 2012.). What is also unusual in HAV is that the empty capsids are extremely robust, whereas for most picornaviruses empty capsids and VLPs tend to be significantly less stable, posing a challenge to the design of VLP-based vaccines. It is possible that HAV may contain cryptic clues as to how to achieve increased stability in empty particles.

In many picornaviruses, the stability of the particle is determined by the interface separating the twelve pentameric assemblies comprising the icosahedral capsid (Filman et al. 1989; Warwicker 1992; Porta et al. 2013). To determine whether the stability of HAV might arise from the VP2 domain swap at the pentamer interface, we characterized CrPV capsid stability, since these share the domain swap seen in HAV. We found that the stability profile of CrPV was similar to the enteroviruses, which do not. Therefore, the VP2 domain swap cannot be explained as a strategy to achieve the high physical stability of the HAV capsid.

In the HAV capsid, however, there is particularly tight packing in the interaction region adjacent to the 2-fold axes that separates in the initial stages of uncoating in the enteroviruses (Wang et al. 2012)(Garriga et al. 2012). The complementarity in HAV resembles that seen in poliovirus, where the release of pocket factor likely triggers instability. This is achieved by matching tyrosine side-chains with small residues nestled around the 2-fold axis. This could contribute significantly to high HAV particle stability, whilst the VP2 domain swap could be indicative of a fundamental difference in how HAV uncoats its genome compared to enteroviruses. Apart from these two striking features, the other candidate interface, relating to protomers around the icosahedral 5-fold axes, has similar properties to other picornaviruses.

Like enteroviruses, HAV could transfer its genome across the endosomal membrane to the cytoplasm using a mechanism involving the formation of an “umbilicus”-like structure comprised of the amphipathic N-terminal helix of VP1 and a membrane pore generated by VP4 (Ren et al. 2013; Butan et al. 2014). The shortened, non-myristoylated version of VP4 in HAV has some hydrophobic character consistent with the formation of an amphipathic helix, and it has been reported to rupture membranes by formation of discrete pores (Shukla et al. 2014) as has the VP4 from TrV (Sanchez-Eugenía et al. 2015). While HAV is capable of undergoing transcytosis when complexed with antibody (Dotzauer et al. 2000; Dotzauer et al. 2005; Counihan and Anderson 2016), HAV transcytosis is inefficient and likely involves vesicular transport, not entry of intact HAV particles across the plasma or endosomal membrane directly into the cytoplasm. It's not clear what mechanism if any could provide for that, or what specific host factor (perhaps a protease) could be capable of triggering particle disassembly within the cytoplasm. Thus, despite the seemingly unsurmountable challenge posed by the stability of the HAV capsid, uncoating is likely to follow interactions with a receptor expressed on membrane, as in the enteroviruses.

The interface between the pentameric assemblies in the HAV capsid is strengthened by a combination of factors: the domain swap cross-over and the supplementation of beta-sheets that span the two-fold axis, the enhanced complementarity of the helices at the two-fold axis, and the overall tighter knit of the capsid (which has no holes, as mentioned above) compared with other picornaviruses. This makes it quite mysterious how the particle disassembles. By reference to the *Dicistroviridae*, the structure of empty, heated, TrV particles revealed that the capsids were not expanded nor had any channels by which the RNA could have exited contrasting to the enterovirus heated particles. Internally the first 33 residues of VP1 were disordered but more significantly, the first 76 residues of VP2 comprising the entire domain swap portion were also disordered. It was suggested that RNA release may take place through transient flipping out of pentamer subunits (Sanchez-Eugenía et al. 2016). Although there is no direct evidence to support this model it seems likely that TrV and HAV share a common mechanism, and structural analysis, most likely by cryo-EM, of transient or trapped intermediates should be able to resolve this puzzle.

Enveloped HAV and ESCRT engagement remains a structural puzzle

As noted above HAV is the first picornavirus for which nonlytic cellular egress in extracellular vesicles has been found likely to be important in disease pathogenesis. How this happens remains obscure and the origins of the surrounding membranes have also been poorly understood. Recently, however, McKnight et al. (McKnight et al. 2017) reported a quantitative proteomics analysis of eHAV virions, which demonstrated that capsids are selected as cargo for vesicular export via a highly specific process, and that infectious eHAV virions possess a set of host proteins similar to that of exosomes, enriched in endolysosomal components and missing markers of autophagy, suggesting an exosome-like mechanism of endosomal sorting complex required for transport (ESCRT)-mediated eHAV biogenesis involving endosomal budding. Note that, as for

many other aspects of HAV, this is quite different to the autophagosome-mediated release proposed for enteroviruses. Further studies showed that ESCRT-III proteins, particularly CHMP2A, have a role in eHAV biogenesis, as previously proposed (McKnight et al. 2017). This is in-line with the observation (based in knock down and mutagenesis) that HAV harbours tandem YPX₃L motifs (VP2: Y₁₄₄PHGLL₁₄₉ and Y₁₇₇PVWEL₁₈₂), candidates to bind a component of the ESCRT pathway, ALIX (Feng et al. 2013). This would then allow the virus to engage the ESCRT complex to facilitate release of enveloped particles via exocytosis (Feng et al. 2013). As shown in Fig. 4, the residues implicated are surprisingly mostly buried, apart from E181 and P178, making their role unclear unless major conformational rearrangements render them accessible, or if they are accessible on a precursor particle. Structural analysis of eHAV particles might clarify this.

Does the antigenic surface of HAV comprise essentially a single antigenic site?

Several studies have addressed the antigenic structure of the HAV capsid, largely built around the characterization of a number of neutralizing monoclonal antibodies (mAbs). Until recently, the evidence for residues comprising part of the binding sites of such antibodies was indirect, mainly that viruses harbouring mutations at a certain points escaped neutralization. On this basis, residues S102, V171, A176 and K221 of VP1 and D70, S71, Q74 and 102-121 of VP3 were implicated in neutralizing epitopes (Ping and Lemon 1992). All except K221 of VP1 were proposed to form part of a single antigenic site, however this is plainly nonsensical since these residues are separated by 40-50 Å on the particle surface. In a more recent study, 12 mAbs were generated against HAV particles (Wang et al. 2015), one of which, mAb #11, was found to have both neutralizing activity as well as the ability to bind VP3 in immunoblots, providing evidence of direct physical interaction with VP3. Based on the particle structure, peptides representing all exposed regions on the surface of the capsid were synthesized. When tested for their ability to be bound by mAb #11, peptides representing VP3 residues 67-77 and 208-219 reacted mildly. These VP3 peptide sequences are relatively closely positioned on the surface of the virus, and thus may contribute to an augmented version of the major conformational epitope identified by Ping and Lemon previously (Ping and Lemon 1992).

To identify additional epitopes on the surface of HAV, structure-based predictive methodology (Borley et al. 2013) was applied. The results concurred with those above and suggested that VP2, residues 71 and 198, and VP3 residues 89-96 may comprise additional epitopes. More recently direct evidence has come from a high-resolution cryo-EM complex between HAV and a potent HAV-specific monoclonal antibody, R10 (Wang et al. 2017). This antibody is an extremely potent neutraliser (neut₅₀ is ~2 nM, ~0.3 µg/mL for the antibody and ~3nM, ~0.45 µg/mL for the Fab) and RT-PCR experiments to quantify virus remaining on the cell surface demonstrated that not only could the antibody block attachment but it could actively compete with the receptors on the cell surface (Wang et al. 2017). These experiments show that R10 neutralizes HAV infection by blocking attachment to the host cell. The complex thus reveals not only atomic details of the antibody-binding site (Fig. 5), but also suggests where cell-attachment might occur. Despite binding being extremely tight, only the Fab fragment destabilises the capsid, and then only relatively mildly (reducing the melting temperature by 7 °C), whereas the antibody slightly stabilizes the particle (by 2 °C), consistent with bivalent attachment cross-linking adjacent pentameric building blocks (Wang et al. 2017).

The details of the R10 epitope are shown in Fig. 5b. The heavy chain predominantly binds VP3 and contributes ~60% of the protein-protein interface, whereas the light chain binds VP2 and VP3 and contributes ~40%. The interaction surface comprises 4 of the 6 common R10 complementary determining regions (CDRs): H1 (residues 28–32), H2 (residues 52–57), H3 (residues 100–106), and L1 (residues 30–31). Atypically, the light-chain framework region (L1-2, residues 45–55) provides for additional interactions. VP2

residues 64–71 and VP3 residues 68–78, 143–150, 209, and 246 form the epitope on the surface of the capsid. Altogether, these residues are 87.5% identical and 91.7% conserved across the six primate hepatovirus genotypes (see Chapter 4), suggesting that R10 probably binds strongly to most if not all of these viruses. The affinity of R10 for the capsid is strengthened by 19 hydrogen bonds, which are formed in part by the side-chain of Y48 in the light chain, and main-chain carbonyl groups from V103 (H-chain) and L53 (L-chain). These interact with VP3 K150 and R209, respectively, and indirectly with VP2 R67. In addition, unpublished results (XW) show that a further four neutralising antibodies attach at essentially the same epitope. So how do these compelling structural results stack up against the previous serological results? In brief, a subset of the residues previously mapped as being part of the epitope are confirmed, as is even the proposal, from *in silico* analysis, that VP2 might be part of the epitope. However, there is absolutely no possibility of any of the VP1 residues directly influencing binding. We will return to this puzzle after considering what these results might tell us about receptor binding.

All the above results are in-line with neutralisation being achieved through blocking of receptor attachment. This pin-points the pentamer interface as the site for receptor recognition. Surprisingly, the crystal structure revealed that the surface of the virus shows a marked segregation of charge broadly: the pentamers are edged by a positively charged fringe, with the central region being mainly negatively charged. The positive charges on the fringe are decorated with a string of sulphate ions derived from the crystallization media. In addition, there are 10 sulphates grouped close to each icosahedral five-fold axis. This might suggest that sulphated glycosaminoglycans may act as, or form part of, the receptor and is supported by competition assays where virus is pre-incubated with heparin sulphate which show marked inhibition at micromolar concentration (Wang et al. 2017). These results, together with our observation that the R10 Fab destabilizes the capsid, suggest the use of a receptor mimic mechanism to neutralize virus infection, suggesting new opportunities for therapeutic intervention.

We are still left with the puzzle of HAV apparently possessing a single, hugely extended, antigenic site. There are examples for other viruses where mutations resulting in altered antigenic properties are not always directly involved in binding antibody (Parry et al. 1990). Now since the VP1 residues are only implicated by escape mutation data we propose that the predominate neutralising epitope is that visualised by Wang et al., (Wang et al. 2017), and that the escape observed via changes in VP1 is achieved by some other mechanism. The mechanism remains unknown, but since the virus is very rigid, and many of these neutralising antibodies are quite potent binders, perhaps the most likely explanation is that escape is achieved by circumventing the block on receptor binding for instance by relaxing the need for receptor attachment to the VP2/3 epitope area, conceivably by boosting receptor attachment (perhaps an alternative receptor) to VP1. In this respect, we note that one of the escape mutations (VP1 221) occurs at a residue also implicated in glycoporphin attachment, whilst another, VP1 102, is adjacent to a sulphate binding site, with two others, VP1 171 and 176 being relatively close by.

Concluding Remarks

There are subtle but profound differences in the structure of the HAV capsid compared to that of previously characterized picornaviruses. A structure-based phylogeny (Fig. 3b) suggests that HAV may represent a link between “classical” mammalian picornaviruses and more primitive picorna-like viruses that infect insects. The striking N-terminal domain swap in VP2 renders its structure more similar to the homologous VP1 and VP3 proteins, providing support for the concept that HAV retains structural and functional features characteristic of primordial picornavirus ancestors, which bore stronger hallmarks of their T=3 origins. Such

archaic viruses may have been more akin to present day insect picorna-like viruses, and the enormous diversity that characterizes present day picornaviruses may have been generated by evolutionary pressure to evolve efficient mechanisms of cell entry. Further research into this virus, in particular structural analyses of the enveloped particles, as well as the capture of assembly states and attachment/uncoating intermediates, is expected to explain more of its unique biology and bridge gaps in our knowledge of the virosphere.

Acknowledgements.

This work was supported by the National Major Project of Infectious Disease, the Ministry of Science and Technology 973 Project (grant no. 2014CB542800), National Science Foundation Grant 81330036 and the Strategic Priority Research Program of the Chinese Academy of Sciences, Grant No. XDB08020200. DIS & EEF are supported by the UK Medical Research Council, grant (grant No. G1000099) and JR by the Wellcome Trust.

References

- Acharya R, Fry E, Stuart D, Fox G, Rowlands D, Brown F. 1989. The three-dimensional structure of foot-and-mouth disease virus at 2.9 Å resolution. *Nature* **337**: 709-716.
- Axford D, Owen RL, Aishima J, Foadi J, Morgan AW, Robinson JI, Nettleship JE, Owens RJ, Moraes I, Fry EE et al. 2012. In situ macromolecular crystallography using microbeams. *Acta crystallographica Section D, Biological crystallography* **68**: 592-600.
- Borley DW, Mahapatra M, Paton DJ, Esnouf RM, Stuart DI, Fry EE. 2013. Evaluation and use of in-silico structure-based epitope prediction with foot-and-mouth disease virus. *PloS one* **8**: e61122.
- Butan C, Filman DJ, Hogle JM. 2014. Cryo-electron microscopy reconstruction shows poliovirus 135S particles poised for membrane interaction and RNA release. *Journal of virology* **88**: 1758-1770.
- Chapman MS, Rossmann MG. 1993. Comparison of surface properties of picornaviruses: strategies for hiding the receptor site from immune surveillance. *Virology* **195**: 745-756.
- Counihan NA, Anderson DA. 2016. Specific IgA enhances the transcytosis and excretion of hepatitis A virus. *Sci Rep* **6**: 21855.
- Das A, Hirai-Yuki A, Gonzalez-Lopez O, Rhein B, Moller-Tank S, Brouillette R, Hensley L, Misumi I, Lovell W, Cullen JM et al. 2017. TIM1 (HAVCR1) Is Not Essential for Cellular Entry of Either Quasi-enveloped or Naked Hepatitis A Virions. *mBio* **8**.
- De Colibus L, Wang X, Spyrou JA, Kelly J, Ren J, Grimes J, Puerstinger G, Stonehouse N, Walter TS, Hu Z et al. 2014. More-powerful virus inhibitors from structure-based analysis of HEV71 capsid-binding molecules. *Nature structural & molecular biology* **21**: 282-288.
- Dotzauer A, Brenner M, Gebhardt U, Vallbracht A. 2005. IgA-coated particles of Hepatitis A virus are translocated antivectorially from the apical to the basolateral site of polarized epithelial cells via the polymeric immunoglobulin receptor. *The Journal of general virology* **86**: 2747-2751.
- Dotzauer A, Gebhardt U, Bieback K, Gottke U, Kracke A, Mages J, Lemon SM, Vallbracht A. 2000. Hepatitis A virus-specific immunoglobulin A mediates infection of hepatocytes with hepatitis A virus via the asialoglycoprotein receptor. *Journal of virology* **74**: 10950-10957.
- Feng Z, Hensley L, McKnight KL, Hu F, Madden V, Ping L, Jeong SH, Walker C, Lanford RE, Lemon SM. 2013. A pathogenic picornavirus acquires an envelope by hijacking cellular membranes. *Nature* **496**: 367-371.

- Filman DJ, Syed R, Chow M, Macadam AJ, Minor PD, Hogle JM. 1989. Structural factors that control conformational transitions and serotype specificity in type 3 poliovirus. *The EMBO journal* **8**: 1567-1579.
- Garriga D, Pickl-Herk A, Luque D, Wruss J, Caston JR, Blaas D, Verdaguer N. 2012. Insights into minor group rhinovirus uncoating: the X-ray structure of the HRV2 empty capsid. *PLoS pathogens* **8**: e1002473.
- Kalynych S, Palkova L, Plevka P. 2015. The Structure of Human Parechovirus 1 Reveals an Association of the RNA Genome with the Capsid. *Journal of virology* **90**: 1377-1386.
- Kaplan G, Totsuka A, Thompson P, Akatsuka T, Moritsugu Y, Feinstone SM. 1996. Identification of a surface glycoprotein on African green monkey kidney cells as a receptor for hepatitis A virus. *The EMBO journal* **15**: 4282-4296.
- Liu Y, Eisenberg D. 2002. 3D domain swapping: as domains continue to swap. *Protein science : a publication of the Protein Society* **11**: 1285-1299.
- McKnight KL, Xie L, Gonzalez-Lopez O, Rivera-Serrano EE, Chen X, Lemon SM. 2017. Protein composition of the hepatitis A virus quasi-envelope. *Proceedings of the National Academy of Sciences of the United States of America* **114**: 6587-6592.
- Parry N, Fox G, Rowlands D, Brown F, Fry E, Acharya R, Logan D, Stuart D. 1990. Structural and serological evidence for a novel mechanism of antigenic variation in foot-and-mouth disease virus. *Nature* **347**: 569-572.
- Ping LH, Lemon SM. 1992. Antigenic structure of human hepatitis A virus defined by analysis of escape mutants selected against murine monoclonal antibodies. *Journal of virology* **66**: 2208-2216.
- Plevka P, Perera R, Yap ML, Cardoso J, Kuhn RJ, Rossmann MG. 2013. Structure of human enterovirus 71 in complex with a capsid-binding inhibitor. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 5463-5467.
- Porta C, Kotecha A, Burman A, Jackson T, Ren J, Loureiro S, Jones IM, Fry EE, Stuart DI, Charleston B. 2013. Rational engineering of recombinant picornavirus capsids to produce safe, protective vaccine antigen. *PLoS pathogens* **9**: e1003255.
- Probst C, Jecht M, Gauss-Muller V. 1999. Intrinsic signals for the assembly of hepatitis A virus particles. Role of structural proteins VP4 and 2A. *The Journal of biological chemistry* **274**: 4527-4531.
- Ren J, Wang X, Hu Z, Gao Q, Sun Y, Li X, Porta C, Walter TS, Gilbert RJ, Zhao Y et al. 2013. Picornavirus uncoating intermediate captured in atomic detail. *Nature communications* **4**: 1929.

- Rossmann MG, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht HJ, Johnson JE, Kamer G, Luo M, Mosser AG et al. 1985. Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature* **317**: 145-153.
- Rossmann MG, Johnson JE. 1989. Icosahedral RNA virus structure. *Annual review of biochemistry* **58**: 533-573.
- Sanchez G, Aragonés L, Costafreda MI, Ribes E, Bosch A, Pinto RM. 2004. Capsid region involved in hepatitis A virus binding to glycophorin A of the erythrocyte membrane. *Journal of virology* **78**: 9807-9813.
- Sanchez-Eugenía R, Durana A, Lopez-Marijuan I, Martí GA, Guerin DM. 2016. X-ray structure of Triatoma virus empty capsid: insights into the mechanism of uncoating and RNA release in dicistroviruses. *The Journal of general virology* **97**: 2769-2779.
- Sanchez-Eugenía R, Goikolea J, Gil-Carmona D, Sanchez-Magraner L, Guerin DM. 2015. Triatoma virus recombinant VP4 protein induces membrane permeability through dynamic pores. *Journal of virology* **89**: 4645-4654.
- Shakeel S, Dykeman EC, White SJ, Ora A, Cockburn JJB, Butcher SJ, Stockley PG, Twarock R. 2017. Genomic RNA folding mediates assembly of human parechovirus. *Nature communications* **8**: 5.
- Shukla A, Padhi AK, Gomes J, Banerjee M. 2014. The VP4 peptide of hepatitis A virus ruptures membranes through formation of discrete pores. *Journal of virology* **88**: 12409-12421.
- Siegl G, Frosner GG, Gauss-Müller V, Tratschin JD, Deinhardt F. 1981. The physicochemical properties of infectious hepatitis A virions. *The Journal of general virology* **57**: 331-341.
- Smith TJ, Kremer MJ, Luo M, Vriend G, Arnold E, Kamer G, Rossmann MG, McKinlay MA, Diana GD, Otto MJ. 1986. The site of attachment in human rhinovirus 14 for antiviral agents that inhibit uncoating. *Science* **233**: 1286-1293.
- Squires G, Pous J, Agirre J, Rozas-Dennis GS, Costabel MD, Martí GA, Navaza J, Bressanelli S, Guerin DM, Rey FA. 2013. Structure of the Triatoma virus capsid. *Acta crystallographica Section D, Biological crystallography* **69**: 1026-1037.
- Tate J, Liljas L, Scotti P, Christian P, Lin T, Johnson JE. 1999. The crystal structure of cricket paralysis virus: the first view of a new virus family. *Nature structural biology* **6**: 765-774.
- Walter TS, Ren J, Tuthill TJ, Rowlands DJ, Stuart DI, Fry EE. 2012. A plate-based high-throughput assay for virus stability and vaccine formulation. *Journal of virological methods* **185**: 166-170.

- Wang X, Peng W, Ren J, Hu Z, Xu J, Lou Z, Li X, Yin W, Shen X, Porta C et al. 2012. A sensor-adaptor mechanism for enterovirus uncoating from structures of EV71. *Nature structural & molecular biology* **19**: 424-429.
- Wang X, Ren J, Gao Q, Hu Z, Sun Y, Li X, Rowlands DJ, Yin W, Wang J, Stuart DI et al. 2015. Hepatitis A virus and the origins of picornaviruses. *Nature* **517**: 85-88.
- Wang X, Zhu L, Dang M, Hu Z, Gao Q, Yuan S, Sun Y, Zhang B, Ren J, Kotecha A et al. 2017. Potent neutralization of hepatitis A virus reveals a receptor mimic mechanism and the receptor recognition site. *Proceedings of the National Academy of Sciences of the United States of America* **114**: 770-775.
- Warwicker J. 1992. Model for the differential stabilities of rhinovirus and poliovirus to mild acidic pH, based on electrostatics calculations. *Journal of molecular biology* **223**: 247-257.
- Zhu L, Wang X, Ren J, Porta C, Wenham H, Ekstrom JO, Panjwani A, Knowles NJ, Kotecha A, Siebert CA et al. 2015. Structure of Ljungan virus provides insight into genome packaging of this picornavirus. *Nature communications* **6**: 8316.

Figure legends

Figure 1

- a) Schematic of P1 genome organization. The numbers of amino acids in the polypeptides are shown. The colour scheme is VP1: blue, VP2: green, VP3: red, VP4 yellow, this is used throughout the figures.
- b) Surface overview., The HAV accessible surface. White lines clarify the particle facets.

Figure 2

- a) No canyon depression. Side view of surface of protomer of HAV (grey) overlaid on a ribbon depiction of an EV71 protomer with the proteins colour coded as in 1b. This highlights the loops that are truncated in HAV to flatten out the canyon, which are labelled.
- b) Potential glycoporphin A receptor binding site. A pentamer from the HAV capsid with VP1-3 drawn in cartoon representation showing the location of residues implicated in the glycoporphin A receptor binding site as spheres: magenta VP1 Lys 221 and grey 102-221 VP3.
- c) VP1 extensions. Side by side surface renditions of an HAV pentamer (left) and a Ljungan Virus pentamer (right). The magenta spheres highlight the last VP1 residue visible in HAV beyond which are approximately 20 VP1 residues and the 2A extension that would be present in the enveloped particle. On the Ljungan virus pentamer the residues corresponding to the extended VP1 C-terminus are drawn in royal blue.
- d) Lack of hydrophobic pocket in VP1 beta-barrel compared to an enterovirus. An HAV protomer (left) is drawn in cartoon representation alongside that for EV71 (right). The pocket volume calculated with PyMOL is drawn in light blue mesh. The lipid that binds in the EV71 pocket is drawn in magenta sticks.

Figure 3

- a) VP2 domain swap. A surface rendered protomer of HAV is depicted with the VP2 N-terminus differentiated by being drawn in green cartoon representation. The VP2 N-termini from LV and FMDV are superimposed and depicted in cyan and magenta, respectively. It can be seen that both LV and HAV share the domain swap conformation characteristic of insect picorna-like viruses rather than the classical picornavirus conformation represented by FMDV.
- b) Structure-based phylogenetic tree of representative picornaviruses and cripaviruses. Included in the analysis are 3VBF (EV71), 1BEV (bovine enterovirus), 4HRV (human rhinovirus14), 1EAH (poliovirus type2), 1COV (coxsackievirus B3), 1TME (Theilers virus), 3MEV (mengo virus), 3CJI (Seneca valley virus), 1ZBA (FMDV A10), 2WFF (equine rhinitis A virus), 3NAP (triatoma virus), 5GKA (aichi virus), 3JB4 (Ljungan virus) and 1B35 (CrPV). Evolutionary distance is calculated based on the number of unmatched residues and the deviation among matched residues. Residues corresponding to the HAV VP2 switch region (1-53) are excluded although their inclusion does not affect the result).
- a) Disordered region in the empty particle. HAV is viewed from the inside with blue mesh corresponding to positive $|F_o - F_c|$ electron density calculated taking the correctly positioned empty HAV from the full HAV. This shows that VP1 2-28 and VP2 5-17 around the three-fold axis are better defined in the full particle.

Figure 4

Location of ESCRT late domain motifs. Late domain YPX₃L motifs are shown as magenta spheres on the pentamer of HAV (as drawn in Fig.2A). In the inset box a close up of the residues that are accessible is shown in the context of the accessible surface.

Figure 5

a) Surface of the cryo-EM complex for R10 and HAV based on the 4.2 Å resolution cryo-EM structure of the R10 Fab:HAV full particle complex. The virus surface is colour coded green to yellow according to distance from the particle centre whilst the Fab molecules are shown in red. One icosahedral asymmetric unit is indicated by a white triangle. b) Surface maps of HAV generated using RIVEM (Chapman and Rossmann 1993) Residues of VP1, VP2, and VP3 are shaded in pale blue, green and red, respectively; residues involved in binding to R10 are shown in brighter colors corresponding to the protein chain they belong to, and are outlined by blue and yellow lines. An icosahedral protomer is outlined in purple with the symmetry axes labeled. Panel (a) is reproduced from Wang et al. 2017, and panel (b) is modified from the same paper.