**Virology**

The structure of enteric human adenovirus 41—A leading cause of diarrhea in children

K. Rafie1,2,3*, A. Lenman4,5,*, J. Fuchs6, A. Rajan4,7, N. Arnbarg4†, L.-A. Carlson1,2,3†

Human adenovirus (HAdV) types F40 and F41 are a prominent cause of diarrhea and diarrhea-associated mortality in young children worldwide. These enteric HAdVs differ notably in tissue tropism and pathogenicity from respiratory and ocular adenoviruses, but the structural basis for this divergence has been unknown. Here, we present the first structure of an enteric HAdV—HAdV-F41—determined by cryo–electron microscopy to a resolution of 3.8 Å. The structure reveals extensive alterations to the virion exterior as compared to nonenteric HAdVs, including a unique arrangement of capsid protein IX. The structure also provides new insights into conserved aspects of HAdV architecture such as a proposed location of core protein V, which links the viral DNA to the capsid, and assembly-induced conformational changes in the penton base protein. Our findings provide the structural basis for adaptation of enteric HAdVs to a fundamentally different tissue tropism.

**Introduction**

Adenoviruses (AdVs) are common pathogens in human, causing diseases not only in airways, eyes, and intestine but also in the liver, urinary tract, and/or adenoids (1). To date, more than 100 human AdV (HAdV) types have been isolated, characterized, and classified into seven species (A to G) (2). In addition to causing serious diseases in humans, several AdVs are being explored as vaccine vehicles against infectious diseases such as coronavirus disease 2019, Middle East respiratory syndrome (MERS), Ebola disease, AIDS, Lassa fever, and Zika disease (3). These efforts include a vaccine candidate based on HAdV-F41 that elicits neutralizing antibodies against MERS coronavirus in vivo (4). The two sole members of HAdV species F, HAdV-F40 and HAdV-F41, stand out as the only HAdVs with a pronounced gastrointestinal tropism. These so-called enteric AdVs are a leading cause of diarrhea and diarrhea-associated mortality (5) in young children, inferior only to Shigella and rotavirus (6). Diarrhea is estimated to cause ~530,000 deaths/year in children younger than 5 years worldwide (7). Thus, there are strong incentives to understand the structural and molecular basis of enteric HAdVs.

AdVs are double-stranded DNA viruses with an ~35–kilo–base pair genome sheltered in a large (~950 Å in diameter), nonenveloped capsid with icosahedral symmetry (8–13). At each of the 12 capsid vertices, penton base (PB) subunits organize as homopentamers (14), anchoring the N-terminal tails of the protruding, trimeric fibers (15, 16) to the capsid. Another so-called major capsid protein is the hexon protein (17), present in 240 trimers per virion. Hexons are the main structural component of the virion facets and are organized to give the virion a pseudo T = 25 symmetry (Fig. 1A). Hexon assemblies are stabilized by minor capsid protein IIIa (IIIa), VI, and VIII (located in the capsid interior) and by IX (exposed on the capsid exterior) (9, 10, 18). To date, two HAdVs, HAdV-C5 (11, 12, 19, 20) and HAdV-D26 (21), as well as individual capsid proteins or their subdomains (14, 17, 22) of multiple AdV types have been structurally determined at high resolution.

The enteric HAdVs have adapted to a distinct tissue tropism from other AdVs, which is presumably reflected in their capsid structure. One major known difference is that enteric HAdVs contain two different types of fiber proteins, long and short (23, 24), whereas other AdVs contain only one type of fiber. Virions dock on to cells through fiber interactions with cellular receptors, followed by internalization mediated by PB interactions with cellular integrins (25). HAdV-F41 long fibers bind to the Coxsackievirus and AdV receptor (26), but no binding partners have been identified for the short fiber. The short fibers—structural hallmarks of enteric HAdVs—contribute to resistance to low pH of enteric HAdVs (27). All other HAdVs contain a conserved, integrin-interacting Arg-Gly-Asp (RGD) motif in the PB (28). Notably, the enteric HAdV-F40 and HAdV-F41 lack this conserved RGD motif and thus use different integrins for entry (29), which may explain their different and much slower entry mechanism (30, 31). Knowledge about the capsid proteins, their structural organization, and host molecule interactions is important for design and development of AdVs as vectors and vaccine vehicles.

Despite the medical importance of enteric AdVs as a major cause of childhood mortality through diarrhea, the structural basis for their infection is not known. We used cryo–electron microscopy (cryo-EM) to determine near–atomic structures of the HAdV-F41 virion at pH 7.4 and at pH 4.0, the latter sets as an average of the diurnal pH in the stomach of young children (32). These structures reveal a capsid that is structurally unchanged by stomach pH and has an extensively remodeled surface as compared to nonenteric HAdVs. We further propose a conserved location of core protein V, which links the AdV genome to the capsid. Last, we describe the assembly-induced structural changes to the PB protein. We believe that these findings will lay the foundation for a detailed molecular understanding of enteric AdVs, how to prevent their infection, and how to further explore AdVs as vehicles for vaccine development.
RESULTS

The structure of HAdV-F41 reveals a capsid with an altered surface charge distribution, structurally unaffected by low pH

To elucidate the structural basis of enteric AdV infection, we determined the structure of HAdV-F41 using cryo-EM. In parallel, the genome of the purified virus (strain "Tak") was sequenced, revealing one nonsynonymous mutation (Val77Ala in VIII) compared to the deposited sequence for the same strain (GenBank: DQ315364.2). Further, its proteome was determined using the high-recovery filter-aided sample preparation (FASP) mass spectrometry (MS) (33), revealing a total of 22 viral proteins present in the purified virus (table S1). At an average resolution of 3.8 Å, the three-dimensional (3D) reconstruction of HAdV-F41 had continuous electron density with well-defined secondary structure elements and side-chain density (fig. S1 and movie S1). Local resolution estimates revealed a resolution of better than 3 Å for large parts of the icosahedral capsid (fig. S1). This allowed us to build and refine an atomic model of the asymmetric unit (ASU), which describes the icosahedral part of the virus (Fig. 1B and movie S2). The final ASU model contained four hexon homotrimeric, single chains of the PB protein and IIIa (fig. S2), 10 chains of VI, two chains of VIII (fig. S2), four chains of the triskelion protein IX, and five chains of unknown identity, confirming known and revealing unknown protein-protein interaction surfaces (table S2). Electron density for the fibers was present at the interface with the icosahedral capsid but was of insufficient quality for extensive model building due to increasing flexibility in more distal parts. Compared to the two other reported structures of HAdVs, HAdV-C5 [Protein Data Bank (PDB): 6B1T (20)] and HAdV-D26 [PDB: 5TX1 (21)], the sequence identity of the capsid proteins ranges from 30 to 80% (table S3). This generally correlated with the average structural similarity between the three structures, with more divergent proteins showing higher structural difference in terms of Cα root mean square deviation (RMSD) (table S3).

In their adaptation to gastrointestinal tropism, a major obstacle for enteric AdVs was likely the passage through the low pH of the stomach to their intestinal site of infection. To investigate the adaptation to the hostile environment of the stomach, we solved the structure of HAdV-F41 at pH 4.0, which resembles the diurnal average gastric pH of young children (Fig. 2A) (32). At the achieved resolution of 5.0 Å, the overall structure of the capsid at pH 4.0 was largely unchanged (overall Cα RMSD of 3.0 Å for proteins listed in Fig. 2B), and no marked local movements were observed (Fig. 2B), showing that the icosahedral part of the HAdV-F41 capsid (which does not include the fibers) does not undergo any large conformational changes at gastric pH. We reasoned that the gastrointestinal adaptation might have altered the distribution of acidic and basic residues exposed on the outer surface of the capsid. To investigate this, the surface

![Fig. 1. The overall structure of HAdV-F41.](http://advances.sciencemag.org/)

(A) Schematic representation of the capsid and core structure of HAdV-F41. (B) Surface representation of the HAdV-F41 electron density with one asymmetric unit (ASU) highlighted in red (left) and a surface representation of the ASU of the HAdV-F41 atomic model viewed from the virion exterior (middle) and interior (right).
charge distribution of HAdV-F41 along with the other two existing HAdV structures (HAdV-C5 and HAdV-D26) was calculated for pH 7.4. A visual comparison of the charge distribution revealed substantial differences between the three HAdVs (Fig. 2C). The exposed surface of HAdV-D26 is almost entirely covered by negative charge at pH 7.4, and HAdV-C5 has mostly negatively charged surfaces on the top of the hexons [in a calculation underestimating the amount of negative charge due to exclusion of the flexible, and highly negatively charged, hypervariable region 1 (HVR1; Fig. 2D), which was not built in any of the HAdV-C5 structures (19, 20, 34–37)]. By comparison, the capsid of HAdV-F41 is predominantly uncharged at pH 7.4, especially at the top of the hexon towers and PB protein (Fig. 2C). The surface charge distribution of HAdV-F41 at pH 4.0 revealed two distinct regions as still being relatively uncharged at this extreme pH: the N-terminal part of IX, largely occluded between hexons, and the solvent-exposed loops at the top of the hexons (Fig. 2C).

Whereas the overall structural fold of hexon chains is conserved among AdVs, they differ in the seven HVRs. Comparing HAdV-F41 (at pH 7.4) to HAdV-C5 and HAdV-D26, substantial differences were found in all seven HVRs (fig. S3 and table S4). In particular, HVR1 stands out in the comparison because it is much shorter in enteric HAdVs (Fig. 2D). The highly negatively charged HVR1 loop has not been built in any reported HAdV-C5 structure (11, 20), indicating its flexibility. On the other hand, the much shorter HVR1 in HAdV-F41 forms a loop with a rigid conformation that allowed tracing the entire length of the polypeptide (Fig. 2E).

Together, these findings show that the icosahedral part of the HAdV-F41 capsid is structurally unperturbed by exposure to stomach-like pH and has evolved to expose fewer charged residues on its exterior as compared to nonenteric HAdVs, most prominently exemplified by a near-complete deletion of the HVR1.

**Protein IX arranges in a unique manner in HAdV-F41**

Among the so-called minor capsid proteins, IX forms the most extended and complex arrangements. In previously reported structures of HAdV-C5 and HAdV-D26, this amounts to a tight mesh of

---

**Fig. 2. Structure and surface charge distribution of HAdV-F41 at pH 7.4 and pH 4.0.** (A) Surface representation of the HAdV-F41 electron density at pH 7.4 and pH 4.0, colored by distance from the virion center. (B) pH-dependent structural changes in selected HAdV-F41 capsid proteins as measured by RMSD at Cα level. (C) Surface charge distribution for the atomic models of HAdV-C5 (PDB: 6CGV), HAdV-D26 (PDB: 5TX1), and HAdV-F41 at pH 7.4 as well as HAdV-F41 at pH 4.0. Red represents a local net negative charge, blue represents positive, and white represents uncharged. (D) Comparison of the HVR1 sequences between HAdV-C5, HAdV-D26, HAdV-F40, and HAdV-F41. (E) Cartoon representation of the HVR1-containing loop for HAdV-C5 and HAdV-F41. The unbuilt loop for HAdV-C5 is shown as a dashed line.
ordered protein density that stretches through the canyons between hexons across the virion surface (Fig. 3A) (11, 21, 38). In HAdV-F41, only the N-terminal residues 1 to 58 (henceforth "IX-N") were sufficiently ordered to trace the protein chain, thus ruling out the same sort of ordered, virus-spanning IX cage seen in other HAdVs (Fig. 3A). To further investigate whether there was any ordered protein density where the C terminus of IX resides in HAdV-C5 and HAdV-D26, we performed a localized asymmetric 3D classification at this position.

Fig. 3. The triskelion-forming protein IX assembles in a unique way in HAdV-F41. (A) Surface representation of IX assembly in HAdV-D26 and HAdV-F41 including all ordered protein density. (B) Graphical representation of HAdV-F41 IX-N triskelion assembly. The three IX chain atoms (green, gray, and purple) are shown in stick representation covered by a semitransparent surface. (C) Close-up view of the hydrophobic core at the center of the IX triskelion assembly, formed by residues Phe17, Phe12, and Tyr20 of each IX chain. (D) Computational slice through the electron density of HAdV-F41 at pH 7.4, at the position of IX at the local threefold axis. Electron density is white except the proposed density of IX-C, which is highlighted in green. (E) As (D), but at the icosahedral threefold axis. (F) Schematic representation of one possible arrangement of IX-C.
Neither of the 3D classes showed any ordered protein density corresponding to the four-helical bundle that IX forms in the other HAdVs at this position (fig. S4). Analogously to other HAdVs, IX-N trimerizes to form a triskelion (Fig. 3B) located between three non-peripentonal hexon units (Fig. 1B). Each facet of the virion harbors four copies of the IX-N triskelion in two distinct structural surroundings: three copies at the local threefold symmetry axes of each ASU and one at the icosahedral threefold symmetry axis at the center of the facet (fig. S5). The conformations of IX-N in these two surroundings are virtually identical (Fig. 3B). The three IX chains come together at the center of this triskelion, where interactions between residues Phe$^{12}$, Phe$^{17}$, and Tyr$^{20}$ from each chain form a hydrophobic core (Fig. 3C). This hydrophobic core is differently arranged compared to the nonenteric HAdV-C5 and HAdV-D26 (fig. S5) and has more large hydrophobic residues at its center.

Sequence comparisons of HAdV-C5, HAdV-D26, and HAdV-F41 protein IX revealed that their respective IX-N have a higher sequence homology than the C-terminal parts of IX (residues 59 to 133; “IX-C”) (fig. S5). However, the sequences of IX-C are near identical in HAdV-F41 and the related HAdV-F40, indicating conservation between enteric AdVs. MS analysis detected the entire IX sequence in the purified HAdV-F41, confirming the presence of IX-C in the purified virions (table S1). We reasoned that the substantial protein mass corresponding to IX-C, emanating in the constricted space between the hexons, should be visible in the electron density map at a lower threshold even if it is flexible. Indeed, the interhexonal space above IX-N harbors electron density corresponding to a flexible protein at the local threefold axis (fig. S4). In a localized asymmetric reconstruction, the three HVR2 could be resolved in their entirety, showing that they are present in a single conformation and form a conformation of defined size (fig. S6) from which this density protrudes. In contrast, there is clearly no electron density above IX-N at the icosahedral threefold position (Fig. 3E and fig. S4), suggesting that the spatial organization of IX-C differs between these two positions (Fig. 3F). Adjacent to the position where the C terminus of IX residues in HAdV-C5 and HAdV-D26, one minor class representing 15% of the positions had a weak density somewhat resembling the protruding density at the local threefold axis (fig. S4).

Together, these data suggest a very different arrangement of IX in enteric AdVs as compared to respiratory (C5) and ocular (D26) HAdVs. In HAdV-F41, the C-terminal half of IX is flexible and appears to expose its C terminus to the capsid exterior at three of the four IX positions in each virion facet.

The HAdV-F41 PB undergoes assembly-induced conformational changes

Located on the fivefold symmetry axes of AdV capsids, the PB protein forms a homopentamer that contains integrin-binding motifs and serves as an assembly hub connecting the icosahedral capsid to the fibers (Fig. 1A). In our reconstruction of the entire HAdV-F41, the electron density for the PB was less well resolved than other parts of the capsid. To improve the map of the PB, we performed a localized asymmetric reconstruction of the PB monomer (fig. S7). The improved map allowed the building of an atomic model for the PB, which was placed into a composite atomic model of the entire ASU. Overall, the HAdV-F41 PB is very similar to the PB in HAdV-C5 (11, 12, 20) and HAdV-D26 (21), with a β sheet–rich fold that can roughly be divided into four domains: crown, head, body, and tail, with the body and head as the main domains, separated by loop regions (Fig. 4A).

During assembly of the virus particle, the PB forms a plethora of interactions with peripentonal hexons, the fiber, and IIIa (table S2). To investigate conformational changes induced during the PB assembly process, we solved the structure of a recombinantly expressed HAdV-F41 PB in solution [free PB (fPB)] by cryo-EM. The map had an average resolution of 3.7 Å (fig. S8), allowing for the placement of an atomic model (Fig. 4B). Comparing the atomic models of the fPB and the virion-bound PB (vPB), the overall Cα shift (RMSD) was very small (~0.9 Å). However, color-coding vPB by its structural deviations from fPB revealed regions with higher RMSD, indicating localized assembly-induced conformational changes (Fig. 4C and movie S3). Moreover, four sequence segments that were built in the vPB model could not be built in the fPB model (Fig. 4A), indicating that these regions are disordered in solution and only become stabilized in a defined conformation upon assembly into the virion. One such region is the tail, a 17-residue (Thr$^{33}$-Gly$^{49}$) random coil region (Fig. 4A), which is disordered in the fPB (Fig. 4B). It is stabilized through interactions with two loop regions from IIIa (table S2). The sequence of the tail domain is largely conserved between HAdVs (fig. S9), suggesting a conserved role as an assembly motif. The second motif (Tyr$^{497}$-Leu$^{505}$) becoming ordered upon assembly is an α helix consisting of residues Gln$^{416}$-Thr$^{427}$, located close to the fivefold axis of the PB (Fig. 4D) and close to where the fiber binds. Although the HAdV-F41 capsid map shows only weak density for the proximal fiber in connection with the PB, the vPB structure did allow tracing of a fragment of the conserved fiber tail (fig. S9 and table S2). Thus, the folding of Gln$^{416}$-Thr$^{427}$ may be dependent on interactions within the capsid and/or binding of the fiber to the PB. Additional assembly-dependent interactions take place in two loop regions located between residues Val$^{70}$-Asn$^{110}$ in the body domain (Fig. 4D). These loops are disordered in the fPB but well resolved in the vPB where their conformation is stabilized by interactions with the peripentonal hexon. The first loop (Ser$^{74}$-Ser$^{79}$), located at the top of the body, is stabilized as an extended coil structure upon interaction with hexon chain Ser$^{663}$-Thr$^{671}$ loop. The second region (Thr$^{100}$-Gln$^{107}$), located at the bottom of body, is stabilized as a short α helix upon binding to an uncharged pocket formed by peripentonal hexon residues Ala$^{1523}$-Ile$^{1540}$.

Peculiarly for enteric HAdVs, the otherwise conserved integrin-binding RGD motif has been replaced by Ile-Gly-Asp-Asp (IGGD) in HAdV-F41 [Arg-Gly-Ala-Asp(RGAD) for HAdV-F40]. In the HAdV-F41 structure, the IGGD-containing loop is the only surface-exposed part of the PB for which we find no continuous electron density (Fig. 4E), despite being shorter than in most other HAdVs (fig. S9) (28). This parallels the observed flexibility of the RGD motifs in the two previously reported structures of HAdVs (19–21), indicating that the function of the IGDG sequence may be dependent on it being flexible until interacting with a target molecule.

In summary, a comparison of the PB in solution and in the virus capsid revealed several distinct motifs that become folded only upon assembly of the PB into the capsid and further revealed that the noncanonical integrin-binding motif IGGD is disordered also in the context of the assembled virus.

The DNA binding protein V is located at a conserved position at the inner face of the capsid

After initial model building of the virion at pH 7.4, the ASU contained five peptide chains that still had not been assigned an identity.
Four of the five chains were deemed too short to assign an identity to. Three of those four chains interlace with different copies of VI at positions where VI interacts with VIII or IIIa at the inner face of the capsid (fig. S10 and movie S2). The fifth unidentified peptide chain is markedly longer and is located at the inner face of the capsid, in a pocket formed by three nonperipentonal hexons (Fig. 5, A and B, and table S5). Its electron density is resolved well enough to identify large side-chain residues, and we thus reasoned that a structural bioinformatics workflow might be devised to reveal its identity. With the initial constraint being only that residues number 4 and 21...
in the identified 24-mer peptide must have large side chains, we used a combination of the proteomics data, exclusion of proteins with known locations, real-space refinement scores, and other considerations to elucidate its identity (see Supplementary Methods and fig. S11 in the Supplementary Materials). After sequential exclusion of candidates based on these criteria, a single most likely candidate remained, a sequence from the center of protein V: Gln170 -Asp 194. V has been reported to bind directly to DNA and to VI, thereby bridging the core and the surrounding capsid, but it has not been localized in any AdV structure. The built sequence of V fits the density without any clashes or unlikely interactions with surrounding proteins and, furthermore, shows a high degree of sequence conservation with V in HAdV-C5 and HAdV-D26 (Fig. 5, C and D, and movie S4). In both hitherto published HAdV structures, HAdV-C5 (20) and HAdV-D26 (21), there is similarly shaped electron density at the corresponding position of the virus capsid (Fig. 5E). In HAdV-C5, no atomic model was built into it (20). Similarly, at the corresponding position in the HAdV-D26 structure, two shorter peptide chains of unknown identity were placed (21). The location of the 24-mer chain of V is such that both the N and C termini of V, which are not resolved in our structure, may protrude toward the interior of the virion in agreement with the proposed role of V to link the viral genome to the capsid. In summary, we used a systematic structural bioinformatics approach to proposing a likely conserved position of core protein V in HAdVs.

DISCUSSION

Here, we present the structure of a major cause of diarrhea and diarrhea-associated mortality in young children: the enteric adenovirus HAdVF41. As the first structure of an AdV with pronounced gastrointestinal tropism, it reveals a capsid whose structure is virtually unaltered by stomach pH and has substantial changes to the virion surface compared to respiratory and ocular HAdVs. Overall, HAdV-F41 has fewer charged, i.e., pH-dependent, residues exposed on the surface of its capsid. This is especially prominent at the top of the hexons where HVR1 is long and rich in negatively charged residues in HAdV-C5. This allows interaction of HAdV-C5 with lactoferrin through a charge-dependent mechanism (40), which contributes to an extended tissue tropism (41, 42). Evolution of HAdV-F41 has resulted in a largely truncated and less charged HVR1 (Fig. 2D), seemingly to adapt to the specific conditions in the gastrointestinal tract. Note that our study could not address one major distinguishing feature of enteric AdVs: the presence of two different fibers (Fig. 1A). The fibers were too flexible to be resolved to a larger extent in the current structure, and it is thus, e.g., possible that the fibers change their conformation at stomach pH more than the icosahedral part of the virus capsid discussed here.

Another major change to the capsid exterior of HAdV-F41 is the starkly different conformation of protein IX. Instead of forming a virus-covering, rigid mesh, the C-terminal half of IX (IX-C) is flexible. An unaccounted density protrudes to the outside of the capsid right above the IX-N triskelia, kept in place by hexon HVR2-containing loops. We favor the interpretation that this density is IX-C due to its proximity to the IX-N triskelia, the presence of IX-C in the purified virus as determined by MS and because IX-C is not found at any other position in the virus structure. Notably, this partial IX-C density is observed above all IX-N trimers except at the icosahedral threefold axis (Fig. 3, D and E, and figs. S4 and S5). In the model, in which this density belongs to IX-C, each IX-C chain could either emerge in cis (i.e., above the same IX-N trimer to which it belongs) or stretch across the virion surface to emerge above another IX-N trimer (in a trans arrangement). The length of IX-C in HAdV-F41 is compatible with both of these arrangements. A cis arrangement of all IX-C would be reminiscent of IX in some non-HAdVs, in which the conformation of IX-C is also more defined (38, 43, 44). Whereas our data do not allow tracing of individual chains of IX-C, the clear lack of any IX-C density above the IX-N trimer at the icosahedral threefold rules out such a pure cis arrangement of IX-C. One possible, parsimonious interpretation would be that the central IX trimer adopts a trans arrangement, donating one IX-C chain to each of the three IX trimers at local threefold positions that, in turn, have their IX-Cs in cis (Fig. 3F). Although other models for the IX-C arrangement may be consistent with our data, it is clear from the data that IX-C in a unique manner in enteric
AdVs compared to other HAdVs studied to date. All these modifications to the virion surface of HAdV-F41 are likely related to the different set of interactors and different pH range that this virus encounters throughout the gastrointestinal tract. Other gastrointestinal viruses interact with components such as bile (calicivirus) (45) and lipopolysaccharides [poliovirus (46), mouse mammary tumor virus (47)], which is crucial for infection of these viruses. Besides low-pH resistant interactions of HAdV-F41 with gastrointestinal phospholipids (48), little is known about the HAdV-F41: gastrointestinal interactions. Finding these interaction partners, e.g., of the disordered and exposed IX-C region, will yield further insights into the infection cycle and tropism of enteric AdVs.

Our study further unveiled how several motifs in the HAdV-F41 PB are disordered in solution and only adopt a defined conformation upon assembly into the virus capsid, laying out another piece of the still largely unfinished puzzle of AdV assembly (10). The observation that the modified integrin-interacting motif of the HAdV-F41 PB is still disordered in the assembled virus particle highlights the need for structural studies of the interactions with its proposed binding partner, laminin-binding integrins (29).

Biochemical data have defined core protein V as a key protein linking the AdV genome to the capsid (49), but, despite its conserved function in AdVs, it had not been located in the AdV capsid. Here, we propose the point of interaction of V to the interior of the capsid and provide data suggesting that this position, at the junction between three nonperipental hexons, is conserved between HAdVs. Previous biochemical data have not suggested V to interact with the hexons but have instead suggested interactions between V and the minor protein VI (49, 50). These data are not mutually exclusive with our identification of the V anchoring point to the hexons, because most of the V sequence is still unaccounted for in the structure and several copies of VI are found in the vicinity of the anchoring point where they may form additional interactions.

Together, the structure of the enteric AdV HAdV-F41 revealed key conserved aspects of AdV architecture and highly divergent features of enteric AdVs, thus laying the foundation for structure-based approaches to developing effective antivirals, for preventing this prominent cause of diarrhea-associated mortality in young children, and for further development of these structurally divergent AdV types as vaccine vehicles.

**MATERIALS AND METHODS**

**Viruse propagation and purification**

Human A549 cells (gift from A. Kidd) were maintained in Dulbecco's modified Eagle medium (DMEM; Sigma-Aldrich) supplemented with 5% fetal bovine serum (FBS; HyClone, GE Healthcare), 20 mM HEPES (Sigma-Aldrich), and penicillin (20 U/ml) and streptomycin (20 μg/ml) (Gibco).

For HAdV-F41 (strain Tak) propagation, 30 bottles of A549 cells (175 cm², at a 90% confluence) were infected with HAdV-F41 inoculation material (produced in A549 cells) in 5 ml of growth media (1% FBS) for 90 min on a rocking table at 37°C. Thereafter, additional 25 ml of growth media (1% FBS) were added to each flask, and the cells were further incubated at 37°C. Infected cells were harvested after approximately 1 week or when cells displayed clear signs of cytopathic effect. Cells were collected by centrifugation, resuspended in DMEM, and disrupted to release virions by freeze-thawing and by addition of equal volume of Vertrel XF (Sigma-Aldrich). After vigorous resuspension, the cell extract was centrifuged at 3000 rpm for 10 min. The upper phase was transferred onto a discontinuous CsCl gradient [densities: 1.27, 1.32, and 1.37 g/ml, in 20 mM tris-HCl (pH 8.0); Sigma-Aldrich] and centrifuged at 25,000 rpm in a Beckman SW40 rotor for 2.5 hours at 4°C. The virion band was collected and desalted on a NAP column (GE Healthcare) into sterile phosphate-buffered saline (PBS).

**Protein identification by MS**

**Protein digestion**

The samples were split for tryptic and chymotryptic digestion and processed using a modified protocol of FASP (33). Briefly, triethylammonium bicarbonate (TEAB) was added to a final concentration of 50 mM TEAB before reduction using 100 mM dithiothreitol at 35°C for 30 min. The reduced samples were loaded onto 10-kDa molecular weight cutoff of Pall Nanosep centrifugal filters (Sigma-Aldrich), washed with 8 M urea and 1% sodium deoxycholate (SDC), and alkylated with 10 mM methyl methane thiosulfonate. Two-step digestion was performed on filters using trypsin and chymotrypsin as digestive enzymes in 50 mM TEAB and 0.5% SDC buffer. The first step was performed overnight, and the second step, with an additional portion of proteases, was performed for 4 hours the next day. Tryptic digestion was performed at 37°C using Pierce MS Grade Trypsin Protease (Thermo Fisher Scientific). Chymotryptic digestion was performed at room temperature using Pierce MS Grade Chymotrypsin Protease (Thermo Fisher Scientific). The peptides were collected by centrifugation, and SDC was precipitated by acidifying the sample with trifluoroacetic acid (final concentration, 1%). The digested sample was desalted using Pierce Peptide Desalting Spin Columns (Thermo Fisher Scientific) according to the manufacturer’s protocol.

**Liquid chromatography–tandem mass spectrometry**

The digested and desalted samples were analyzed using a QExactive HF mass spectrometer interfaced with an Easy-nLC 1200 liquid chromatography–tandem mass spectrometry system (both Thermo Fisher Scientific). Peptides were trapped on an Acclaim PepMap 100 C18 trap column (100 μm by 2 cm; particle size, 5 μm; Thermo Fisher Scientific) and separated on an in-house packed analytical column (75 μm by 300 mm; particle size, 3 μm; Reprosil-Pur C18, Dr. Maisch). A stepped gradient used was from 7 to 35% solvent B in 97 min, followed by an increase to 48% in 8 min and to 100% solvent B in 5 min at a flowrate of 300 nl/min. Solvent A was 0.2% formic acid, and solvent B was 80% acetonitrile in 0.2% formic acid. The mass spectrometer was operated in data-dependent acquisition (DDA) mode where the MS1 scans were acquired at a resolution of 60,000 and a scan range from 400 to 1600 mass/charge ratio (m/z). The 10 most intense ions with a charge state of 2 to 4 were isolated with an isolation window of 1.2 m/z and fragmented using normalized collision energy of 28. The MS2 scans were acquired at a resolution of 30,000, and the dynamic exclusion time was set to 20 s.

**Database search**

Data analysis was performed using Proteome Discoverer (version 1.4, Thermo Fisher Scientific). The data were searched against an in-house database containing the amino acid sequences of HAdV-F41. Mascot (version 2.5.1, Matrix Science) was used as search engine with a precursor mass tolerance of 5 parts per million for MS1 and 30 milli mass unit (mmu) for MS2 spectra. Tryptic peptides were accepted with a maximum of one missed cleavage, and chymotryptic peptides were accepted with maximum three missed cleavages. Variable modification
of methionine oxidation and fixed methylthio of cysteines were selected. The Mascot significance threshold for peptides was set to 0.01.

Cryo-EM sample preparation

Purified HAdV-F41 was used at 4.2 mg/ml (pH 7.4) and 1.6 mg/ml (pH 4.0). The recombinant HAdV-F41 PB (rPB) was purified as described before (29) and used at 1 mg/ml in PBS buffer, supplemented with 5% glycerol. A HAdV-F41 sample at pH 4.0 was prepared by adding 2 μl of a 0.5 M citric acid/1 M Na2HPO4 (pH 3.4) solution to 25 μl of HAdV-F41 (pH 7.4), followed by incubating on ice for 15 min. Samples were vitrified on QUANTIFOIL Cu R200 2/2 (Electron Microscopy Sciences, catalog no. Q2100CR2) and QUANTIFOIL Cu R200 1.2/1.3 (Electron Microscopy Sciences, catalog no. Q3100CR1.3) grids for the virus particles and the recombinant protein, respectively. Before sample application, the grids were glow discharged using a PELOCO easiGlow device (Ted Pella Inc.) at 15 mA for 30 s. Sample was applied by transferring 3 μl of sample onto the glow-discharged side of the grid, blotted, and plunge-frozen in liquid ethane, using a Vitrobot plunge freezer (Thermo Fisher Scientific), with the following settings: 22°C, 80% humidity, a blot force of ~20, and a blotting time of 3 s. For HAdV-F41 at pH 7.4, sample was applied twice with a blotting step, using the same settings as above, between applications (51).

Data collection

All data were collected on an FEI Titan Krios transmission electron microscope (Thermo Fisher Scientific) operated at 300 keV and equipped with a Gatan BioQuantum energy filter and a K2 direct electron detector. A condenser aperture of 70 μm (HAdV-F41 at pH 7.4 and pH 4.0) and an objective aperture of 100 μm were chosen for data collection. A C2 aperture of 100 μm was selected for the PB41 data collection. Coma-free alignment was performed with AutoCtfeSherpa. Data were acquired in parallel illumination mode using EPU (Thermo Fisher Scientific) software at a nominal magnification of 130,000× (1.041-Å pixel size). Both datasets for the HAdV-F41 structure at pH 7.4 were collected in superresolution mode. Because of a preferred orientation of PB41, a second dataset was collected at a 30° tilted stage. Data collection parameters are listed in table S6.

Data processing and structure determination

HAdV-F41 at pH 7.4

Two datasets were collected on HAdV-F41 at pH 7.4 and initially processed independently. Data were initially processed using RELION 3.0 (52) and continued in RELION 3.1 (beta) (53). Beam-induced motion was corrected using RELION’s MotionCor2 (54) implementation, which step the superresolution movies were binned once, and the per-micrograph contrast transfer function (CTF) estimated using GCTF software (55) for all datasets. Particles were manually picked and subjected to reference-free 2D classification, and well-resolved classes were combined and subjected to 3D classification, applying icosahedral symmetry [I3 according to Crowther (36)] and a mask of the capsid structures. A low-pass filtered (50 Å) volume of HAdV-C5 [EMD-7034 (20)] was used as a reference volume. Particles were classified into two classes, resulting in 99% of particles allocated to one well-resolved class that was used for downstream processing. 3D refinement was performed using the output of the 3D classification as a reference model, low-pass filtered to 50 Å, with no additional Fourier padding. Following refinement, data were postprocessed, and the particles were subjected to per-particle CTF refinement, Bayesian polishing, and another round of per-particle CTF refinement. The particles were subjected to an additional round of 3D refinement before combining both datasets and performing a final 3D refinement, with no additional Fourier padding. The resolution was calculated using the gold standard Fourier shell correlation [FSC threshold, 0.143)] to 3.84 Å after postprocessing. Last, the data were corrected for the Ewald’s sphere curvature using RELION, which led to a local improvement of the electron density map with a new average resolution of 3.77 Å. Local resolution estimates were calculated using ResMap (57).

A homology model was generated using the SWISS-MODEL server (58), the HAdV-F41 capsid protein sequences, for which homologs have been structurally determined. The resulting homology model was based on the reported HAdV-D26 structure [PDB: 5TX1 (21)]. The model was manually docked into the HAdV-F41 electron density in ChimeraX (59) and the map corresponding to the ASU extracted. The ASU map was locally sharpened using Phenix’s (60) autosharpen tool. Subsequently, the HAdV-F41 homology model was docked and subjected to an initial round of real-space refinement using Phenix. The structure was fully refined using iterative cycles of Phenix’s real-space refinement and model building in Coot (61).

Localized asymmetric reconstruction

To improve the map quality surrounding the PB monomer, the HVR2-loop containing region (local threefold axis), the icosahedral threefold axis, and the region analogous to the region where the four-helical IX-C bundles in HAdV-C5 and HAdV-D26 are located, the map was improved using the localized asymmetric reconstruction workflow reported by Ilca et al. (39) and implemented in Scipion v2.0 (62). Coordinates for the subparticles were determined in ChimeraX and subsequently located by applying icosahedral symmetry and extracted in Scipion v2.0. Subparticles were subsequently filtered to exclude particles not present within a [−20°, 20°] range from the image plane. The resulting subparticles were then subjected to an asymmetric 3D classification. To increase the probability of convergence during classification, changes in the origins and orientations were not allowed. A subsequent 3D refinement yielded a 3D reconstruction of the PB monomer and the HVR2-loop containing region to a resolution of 3.0 and 3.35 Å, respectively. Average resolutions were calculated according to the gold standard FSC calculations (threshold, 0.143). Data processing statistics for the HVR2-containing loop region and the PB monomer are given in table 7. 3DFSC curves were calculated using the Remote 3DFSC Processing Server (63).

Image processing and model building for HAdV-F41 at pH 4.0

Data were processed as described for the HAdV-F41 at pH 7.4 structure up until the first 3D refinement. The volume HAdV-F41 at pH 7.4 was low-pass filtered to 10 Å and used as a reference. The resolution was estimated to 5.0 Å using the gold standard FSC (threshold, 0.143) after postprocessing. Local resolution estimates were calculated using ResMap. The HAdV-F41 (pH 7.4) model was fitted into the reconstructed HAdV-F41 (pH 4.0) density using ChimeraX, an ASU was extracted, and the resulting map was locally sharpened using Phenix. The model was then further fitted and energy minimized using Namdinator (64).

Data processing and structure determination recombinant HAdV-F41 PB

The HAdV-F41 PB (PB41) data (untitled and tilted at 30°) were processed using RELION 3.1 (beta), with beam-induced motion correction...
Calculation of surface charge distribution
Surface charges for HAdV-C5 [PDB: 6CGV (19)], HAdV-D26 [PDB: 5TX1 (21)], and HAdV-F41 were calculated using the PDB2PQR-APBS software package (67) at pH 7.4 and pH 4.0.

Bioinformatics workflow to determine the protein identity of the unknown chain
For each direction of the unknown chain, a 24-mer poly-alanine chain was manually placed into the respective density and initially real-space refined using Coot. A list of sequences was screened using a job pipeline including a mutation step in Coot and real-space refinement in Phenix. Custom bash scripts written for this purpose are available upon request. An extended description is given in Supplementary Methods.

Molecular graphics and visualization
Figures of protein structures and electron densities were generated using ChimeraX.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/2/eabe0974/DC1

View/request a protocol for this paper from Bio-protocol

REFERENCES AND NOTES
1. W. S. M. Wold, M. G. Ison, Adenoviruses, in Fields Virology, D. M. Knipe, P. M. Howley, Eds. (Lippincott Williams & Wilkins, 2016), vol. 2, chap. 56, pp. 1732–1767.
2. HAdV Working Group, (2019); http://hadvwg.gmu.edu/
3. M. Garofalo, M. Staniszewska, S. Salmaso, P. Caliceti, K. W. Pancer, M. Wieczorek, L. Kuryk, Prospects of replication-deficient adenovirus based vaccine development against SARS-CoV-2. Vaccines 8, 293 (2020).
4. X. Guo, Y. Deng, H. Chen, J. Lan, W. Wang, X. Zou, T. Huang, Z. Lu, W. Tan, Systemic and mucosal immunity in mice elicited by a single immunization with human adenovirus type 5 or 41 vector-based vaccines carrying the spike protein of Middle East respiratory syndrome coronavirus. Immunology 145, 476–484 (2015).
5. G. B. D. D. Collaborators, Estimates of the global, regional, and national morbidity, mortality, and aetiology of diarrhoea in 195 countries: A systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect. Dis. 18, 1211–1228 (2018).
6. L. I. J. A. Platts-Mills, J. Juma, F. Kabir, J. Njace, C. Okoi, D. I. Opersario, J. Uddin, S. Ahmed, P. L. Alonso, M. Antonio, S. M. Becker, W. C. Blackwelder, R. F. Breiman, A. S. Faruque, B. Fields, J. Gratz, R. Haque, A. Hossain, M. J. Hossain, S. Jarju, F. Qamar, N. T. Iqbal, B. Kwambana, I. Mandomando, T. L. McCurry, C. Ochieng, J. B. Ochieng, M. Ochieng, C. Onyango, S. Panchalingam, A. Kalam, F. Aziz, S. Qureshi, T. Ramamurthy, J. H. Roberts, D. Saha, S. O. Sow, S. E. Stroup, D. Suri, B. Tamboura, M. Taniuchi, S. M. Tennant, D. Toema, Y. Wu, A. Zaidi, J. P. Nataro, K. L. Kotloff, M. M. Levine, E. R. Houpt, Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: A reanalysis of the GEMS case-control study. Lancet 388, 1291–1301 (2016).
7. GBD 2017 Diarrhoeal Disease Collaborators. Quantifying risks and interventions that have affected the burden of diarrhoea among children younger than 5 years: An analysis of the Global Burden of Disease Study 2017. Lancet Infect. Dis. 20, 37–59 (2020).
8. S. J. Flint, G. R. Nemerow, Adenovirus composition, structure, and biophysical properties, in Human Adenoviruses: From Villains to Vectors (World Scientific, 2017), chap. 2, pp. 15–41.
9. G. R. Nemerow, P. L. Stewart, V. S. Reddy, Structure of human adenovirus. Curr. Opin. Virol. 2, 115–121 (2012).
10. C. San Martin, Latest insights on adenovirus structure and assembly. Viruses 4, 847–877 (2012).
11. H. Liu, J. Jin, S. B. Koh, I. Atanasov, S. Schein, L. Wu, Z. H. Zhou, Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. Science 329, 1038–1043 (2010).
12. V. S. Reddy, S. K. Natchiar, P. L. Stewart, G. R. Nemerow, Crystal structure of human adenovirus at 3.5 Å resolution. Science 329, 1071–1075 (2010).
13. P. L. Stewart, R. M. Burnett, M. Cyrlaff, S. D. Fuller, Image reconstruction reveals the complex molecular organization of adenovirus. Cell 67, 145–154 (1991).
14. C. Zubieta, G. Schoehn, J. Chroboczek, S. Cusack, The structure of the human adenovirus 2 penton. Mol. Cell 17, 121–135 (2005).
15. M. J. van Raaij, A. Mitra, G. Lavigne, S. Cusack, A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. Nature 401, 935–938 (1999).
16. D. Xia, J. J. Henry, R. D. Gerard, J. Deisenhofer, Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution. Structure 2, 1259–1270 (1994).
17. M. M. Roberts, J. L. White, M. G. Grütter, R. M. Burnett, Three-dimensional structure of the adenovirus major coat protein hexon. Science 232, 1148–1151 (1986).
18. M. Benevento, S. Di Palma, J. Snijder, C. L. Moyer, V. S. Reddy, G. R. Nemerow, A. J. Heck, Adenovirus composition, proteolysis, and disassembly studied by in-depth qualitative and quantitative proteomics. J. Biol. Chem. 289, 11421–11430 (2014).
19. S. Kundhavai Natchiar, S. Venkataraman, T. M. Mullen, G. R. Nemerow, V. S. Reddy, Revised crystal structure of human adenovirus reveals the limits on protein IX quasi-equivalence and on analyzing large macromolecular complexes. J. Mol. Biol. 430, 4132–4141 (2018).
20. X. Dai, L. Wu, R. Sun, Z. H. Zhou, Atomic structures of minor proteins Vn and Vt in human adenovirus. J. Virol. 91, e00850-17 (2017).
21. X. Yu, D. Veesler, M. G. Campbell, M. E. Barry, J. F. Asturias, M. A. Barry, V. S. Reddy, Cryo-EM structure of human adenovirus D26 reveals the conservation of structural organization among human adenoviruses. Sci. Adv. 3, e1602670 (2017).
22. E. Seiradake, S. Cusack, Crystal structure of enteric adenovirus serotype 41 short fiber head. J. Virol. 79, 14088–14094 (2005).
23. A. H. Kiddy, J. Chroboczek, S. Cusack, R. W. Ruigrok, Adenovirus type 40 virions contain two distinct fibers. Virology 192, 73–84 (1993).
24. H. Y. Yeh, N. Pieniazek, D. Pieniazek, H. Gelderblom, R. B. Luftig. Human adenovirus type 41 contains two fibers. Virus Res. 33, 179–198 (1994).
25. T. J. Wickham, P. Mathias, D. A. Cheshire, G. R. Nemerow, Integrons vj3 and vj5 promote adenovirus internalization but not virus attachment. Cell 73, 309–319 (1993).
26. P. W. Roelvink, A. Linzanova, J. G. Lee, Y. Li, J. M. Bergelson, R. W. Fincher, D. E. Brough, I. Kovesdi, T. J. Wickham, The coxsackievirus-adenovirus receptor protein can function as a cellular attachment protein for adenovirus serotypes from subgroups A, C, D, and F. J. Virol. 72, 7909–7915 (1998).
27. E. Rodriguez, C. Romero, A. Rio, M. Miralles, A. Raventos, L. Planells, J. F. Burguero, H. Hamada, J. C. Perales, A. Bosch, M. A. Gassull, E. Fernandez, M. Chillon, Short-fiber
and purified virus particles. A.R. purified recombinant PB protein. K.R. collected cryo-EM data and performed image processing, model building, and validation. K.R., A.L., N.A., and L.-A.C. interpreted structural data. J.F. performed proteomics analysis. K.R., A.L., N.A., and L.-A.C. wrote the original manuscript with input from J.F. All authors reviewed and approved the final manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The scripts used for the bioinformatics analysis are available upon request. Coordinates reported in this study have been deposited with the PDB with accession codes 6Z7N (HAdV-F41 ASU) and 6Z7Q (HAdV-F41 (free) PB). Electron microscopy maps and half maps have been deposited in the Electron Microscopy Data Bank with the accession codes EMD-11108 (HAdV-F41 at pH 7.4), EMD-11111 (HAdV-F41 at pH 4.0), EMD-11112 (HAdV-F41 (free) PB), EMD-11109 (localized asymmetric reconstruction of the HAdV-F41 PB), and EMD-11110 (localized asymmetric reconstruction of the HAdV-F41 HVR2-containing loop).

Submitted 3 August 2020
Accepted 17 November 2020
Published 8 January 2021
10.1126/sciadv.abe0974

**Citation:** K. Rafie, A. Lenman, J. Fuchs, A. Rajan, N. Arnberg, L.-A. Carlson, The structure of enteric human adenovirus 41—A leading cause of diarrhea in children. *Sci. Adv.* 7, eabe0974 (2021).
