AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

STRUCTURE AND ASSEMBLY

Adeno-Associated Virus Receptor-Binding: Flexible Domains and Alternative Conformations through Cryo-Electron Tomography of Adeno-Associated Virus 2 (AAV2) and AAV5 Complexes

Guiqing Hu,[a] Mark A. Silveria,[b] Grant M. Zane,[b] Michael S. Chapman,[b] Scott M. Stagg[a,c]

[a]Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida, USA
[b]Department of Biochemistry, University of Missouri, Columbia, Missouri, USA
[c]Department of Biological Sciences, Florida State University, Tallahassee, Florida, USA

ABSTRACT Recombinant forms of adeno-associated virus (rAAV) are vectors of choice in the development of treatments for a number of genetic dispositions. Greater understanding of AAV's molecular virology is needed to underpin needed improvements in efficiency and specificity. Recent advances have included identification of a near-universal entry receptor, AAVR, and structures detected by cryo-electron microscopy (EM) single particle analysis (SPA) that revealed, at high resolution, only the domains of AAVR most tightly bound to AAV. Here, cryogenic electron tomography (cryo-ET) is applied to reveal the neighboring domains of the flexible receptor. For AAV5, where the PKD1 domain is bound strongly, PKD2 is seen in three configurations extending away from the virus. AAV2 binds tightly to the PKD2 domain at a distinct site, and cryo-ET now reveals four configurations of PKD1, all different from that seen in AAV5. The AAV2 receptor complex also shows unmodeled features on the inner surface that appear to be an equilibrium alternate configuration. Other AAV structures start near the 5-fold axis, but now β-strand A is the minor conformer and, for the major conformer, partially ordered N termini near the 2-fold axis join the canonical capsid jellyroll fold at the βA-βB turn. The addition of cryo-ET is revealing unappreciated complexity that is likely relevant to viral entry and to the development of improved gene therapy vectors.

IMPORTANCE With 150 clinical trials for 30 diseases under way, AAV is a leading gene therapy vector. Immunotoxicity at high doses used to overcome inefficient transduction has occasionally proven fatal and highlighted gaps in fundamental virology. AAV enters cells, interacting through distinct sites with different domains of the AAVR receptor, according to AAV clade. Single domains are resolved in structures by cryogenic electron microscopy. Here, the adjoining domains are revealed by cryo-electron tomography of AAV2 and AAV5 complexes. They are in flexible configurations interacting minimally with AAV, despite measurable dependence of AAV2 transduction on both domains.

KEYWORDS AAV, AAVR, cryo-EM, cryo-ET, structure

Adeno-associated virus (AAV) is a small 25-nm T=1 icosahedral virus with a protein shell encapsidating a single-stranded DNA genome (1, 2). AAV is so named because it was discovered during adenovirus preparations and its replication depends on coinfection with adenovirus or one of several other "helper" viruses, not because there is any structural relation (3–5). AAVs were long regarded as nonpathogenic, an initial rationale for their development as transducing vectors for *in vivo* (and *ex vivo*) gene therapy. (6–9). The recent discovery of AAV sequences inserted into proto-oncogenes of patients with hepatocellular carcinoma (HCC) has prompted vigorous debate about causal links to natural infection and

Editor Colin R. Parrish, Cornell University

Address correspondence to Michael S. Chapman, chapmanms@missouri.edu, or Scott M. Stagg, sstagg@fsu.edu.

The authors declare no conflict of interest.

Received 15 January 2022
Accepted 29 April 2022
Published 8 June 2022

future vector use (10–16). A prevalent view is emerging that there may be a concern for individuals with chronic liver disease (17–19).

Nonetheless, it is an exciting time for gene therapy. After many years in development, the first two *in vivo* treatments have been approved by the U.S. Food and Drug Administration (FDA), using AAV2 and AAV9 vectors, respectively. Luxturna is a treatment for an inherited blindness, and Zolgensma is for spinal muscular atrophy (20, 21). AAV vectors are being used for >150 ongoing clinical trials (https://clinicaltrials.gov/) (22), but challenges await in generalization of the early successes. Deaths in a myotubular myopathy trial likely resulted from immune-toxicity of the high doses needed to achieve therapeutic expression levels with an inefficient transducing vector (23, 24). Doses, measured in vector genomes per body mass, have been well tolerated at up to 1×10^{14} vector genomes (vg)/kg, with all three fatalities occurring at 3×10^{14} vg/kg (23, 25). Structural studies are key to an improved fundamental understanding of AAV's virology and its engineering for vector improvement.

Initial crystallographic structures revealed the 60-fold symmetric part of the capsid. The capsid gene is expressed as three variant viral proteins (VP) due to alternative start codons and splice variants (26). The variants are in-frame, sharing most of their amino acid sequences, and it has become conventional to use common numbering, based on the largest, VP1. Ordered structure becomes visible at about residue 220, or ~20 residues beyond the N terminus of VP3, which constitutes ~80% of the capsid (27). Upstream, VP1 (~10%) and VP2 (~10%) are extended by a common region of 65 usually unseen amino acids that some have proposed to function in nuclear localization (28, 29). Then there is a segment unique to VP1 (VP1u), N-terminal of the VP2 start, that contains a phospholipase A2 (PLA_2) domain that is initially sequestered within the capsid but becomes exposed for endosomal escape on the entry pathway (30–34).

Over 130 variants of human and nonhuman primate AAVs have been identified (35, 36). These are grouped into eight major named and unnamed clades, containing one or more serotypes that are antigenically distinct; i.e., antibodies recognizing one serotype do not cross-react with others (35, 37). The serotypes differ in other properties, such as binding preference to glycan attachment factors and empirically determined tissue tropisms (38, 39). This study uses two representatives, AAV2 and AAV5, as model systems. AAV2 is the type species that is the best characterized. AAV5 is tied for the most distantly related with reference to VP3 amino acid sequence.

This study further characterizes interactions with the near-universal protein receptor, AAVR. AAVR was only recently discovered though unbiased genome-wide screening as a receptor key for entry and trafficking (40). Previously, a serotype-specific variety of glycans had been considered to be "primary receptors," heparan sulfate proteoglycan (HSPG) for AAV2 and sialic acid (SIA) for AAV5 (41–43). However, it has recently been argued that the glycans have less-specific roles than classic receptors, and following virological convention, should be considered attachment factors (44), anchoring viruses to cell surfaces but not mediating productive entry. Several membrane proteins, primarily tyrosine kinase receptors and integrins, were also identified as coreceptors for different serotypes, but they have not figured in several more recent knockout screens (40, 45–54). Current evidence indicates that AAV2 and AAV5 attach to cells using different extracellular glycans, that both viruses depend on AAVR for entry and trafficking, and that then AAV2 (but not AAV5) has a downstream dependence on another host membrane protein, GPR108 (54, 55).

AAVR is a C-terminally anchored transmembrane protein, in which the ectodomain (from the N terminus) consists of a signal peptide, a MANEC domain (motif at N terminus with eight cysteines) then five Ig-like polycystic kidney disease (PKD) domains (56, 57). It is the PKD domains that bind AAV, but surprisingly, there are different serotype-specific domain dependencies (55). For AAV2, PKD2 is most important, but PKD1 has an accessory role, whereas AAV5 is exclusively dependent upon PKD1 (55). These determinations were made by (i) surface plasmon resonance (SPR) measurements using AAV and heterologously expressed AAVR domain fragments, (ii) transduction inhibition through addition of solubilized domain fragments, (iii) knockout through domain-deletion, and (iv) viral overlay assay (40). Concurrent cryogenic-electron microscopy (cryo-EM) structure determinations using

different expressed AAVR fragments, PKD1-5 or PKD1-2, revealed PKD2 bound to AAV2 at 2.8 and 2.4 Å, respectively (44, 58). Even though the samples contained 5- and 2-domain fragments, respectively, only the most tightly interacting domain (PKD2) was revealed. Cryo-electron tomography (cryo-ET) of an N-terminal fusion of maltose-binding protein (MBP) and PKD1-5, combined with cross-linking mass spectrometry (XL-MS) was consistent, showing anchoring of PKD2 to the viral surface, and the PKD3-5 domains emanating radially in at least four configurations (44). Then, in succession, came cryo-EM structures of AAV5 complexes, PKD1-5 at 3.2-Å and PKD12 at 2.5-Å resolution, now showing just the PKD1 domain, which alone had previously been implicated (55, 59, 60). Intriguingly, the homologous PKD1 and PKD2 domains were not accommodated as variations of a single AAVR-binding site on AAV, but were at distinct sites. One could then best imagine evolutionary divergence occurring through an ancestral form that bound both domains, but overlay of the structures eliminated simple explanations with the finding that the domains could not be connected plausibly by the unseen 5-residue linker (60).

Cryo-ET has technical advantages enabling determination of 3D structures of flexible molecules in heterogeneous configurations, such as AAVR with its variable PKD domain orientations. In contrast to single-particle cryo-EM, where a single 2D image from many identical or nearly identical particles (10^4 to 10^6) are aligned and averaged into a 3D reconstruction, in tomography, 3D images of every individual particle are realized by tilting the microscope stage. This technique has some limitations because the sample can only be tilted within a range of angles between $-65°$ and $+65°$. A consequence of this is that the resulting 3D reconstructions have a "missing wedge" of information that can distort the 3D volumes. However, the missing wedge can be filled by averaging between aligned subvolumes containing a structure of interest in different orientations and thus with different missing wedges. A structure can be split into subvolume parts for classification and averaging to characterize variability in heterogeneous regions. This can be a particular advantage for structures such as virus-receptor complexes where different copies of a viral capsid protein could have receptor bound in a different configuration. There have been several successful applications of the approach, to for example, the heterogenous structure of simian immunodeficiency virus (SIV) envelope glycoprotein when bound by CD4 receptor or monoclonal antibody 36D5 (61).

Here, we use cryo-ET to focus on the 2-domain receptor complex of AAV for a holistic and hybrid comparison with single-particle cryo-EM to locate the parts that had been refractory to the high-resolution cryo-EM. It uses the unique advantages of cryo-ET to distinguish different conformational states, focusing reconstructions on the subvolumes surrounding each 3-fold axis to reveal the hitherto unseen domains in the AAV-PKD12 complexes and other elements of both the receptor and virus structures that have been smeared beyond recognition in the 60-fold averaged cryo-EM reconstructions.

RESULTS

The structure of AAV2 bound with PKD1-2. Cryogenic electron tomographic (cryo-ET) tilt series were acquired for AAV2 bound by the PKD1-2 domains of AAVR, and tomograms were reconstructed (Fig. 1A and C). AAV2 subvolumes were aligned and then subdivided into individual trimers, and all aligned trimers were averaged (Figure 2B, part 1). This revealed densities at about 20-Å resolution (Fig. 1C), corresponding to the VP capsid protein and PKD2, consistent with the published single-particle analyses (SPA) of the AAV2/PKD1-2 complexes (44, 58). The EM samples prepared for the previous SPA and the current cryo-ET study were prepared in exactly the same way (44), the SPA yielding 2.4-Å resolution. Side-chain features leave no ambiguity that it is PKD2 that is bound tightly at this site on AAV2 and that the domain orientation is with the N terminus near the viral 2-fold and the C terminus near the viral 3-fold axis. The atomic model of the AAV2 spike fits the global average well, defining the density for the viral protein and the PKD2 domain of AAVR (Figure 2B, part 2). The AAV2 viral protein and the PKD2 domain are readily apparent, but there was no sign of the PKD1 domain. In order to reveal PKD1, tomographic subclassification was performed using a trimer subvolume that extended mostly outward beyond PKD2 for the AAV2/PKD1-2 complex. This revealed

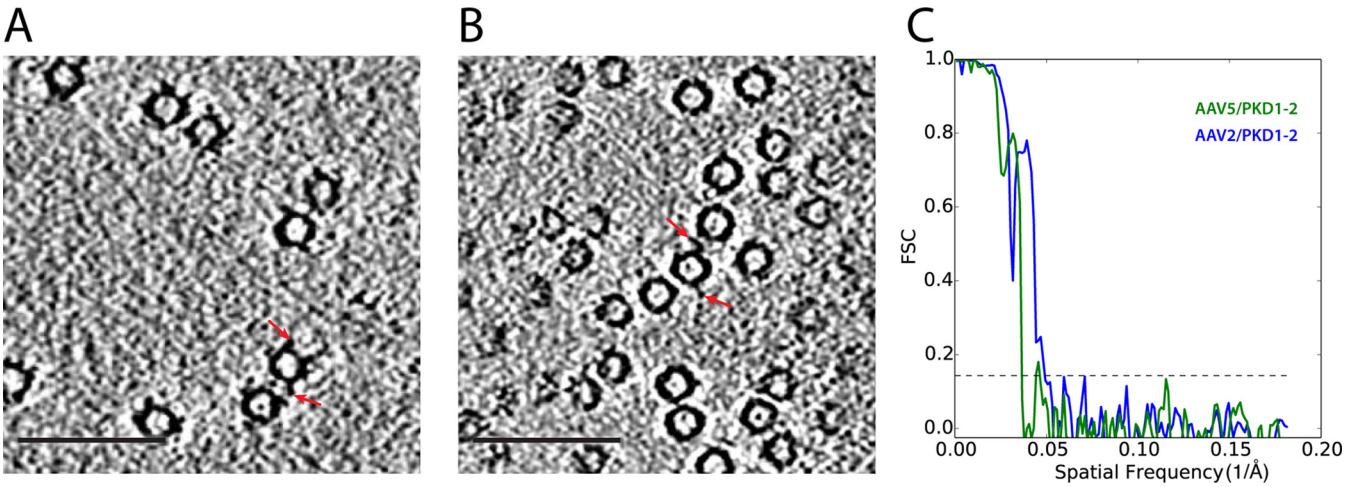

FIG 1 Cryo-ET of AAV-receptor complexes. (A) Slice through a tomogram of AAV2/PKD1-2. AAVR densities are indicated by the red arrows. (B) Slice through a tomogram of AAV5/PKD1-2. (C) Fourier shell correlation (FSC) curves for the subvolume average reconstructions from cryo-ET with estimated resolutions of 20 Å for AAV2/PKD1-2 (blue) and 25 Å for AAV5/PKD1-2 (green). Scale bars: 1,000 Å.

additional features corresponding to four distinct conformations for PKD1 (Fig. 2C). The features are of the correct shape and size for a PKD domain, and an atomic model of PKD1 fits well into the map of each class (Fig. 2D). A 5-residue linker (residues 400 to 404) can be built between the PKD1 and PKD2 domains with plausible stereochemistry. The length of the 5-residue linker constrains interpretation of the observed domain conformers to be PKD1 linked to PKD2 in a head-to-tail configuration, the C-terminal end of PKD1 closer to PKD2 and the viral surface.

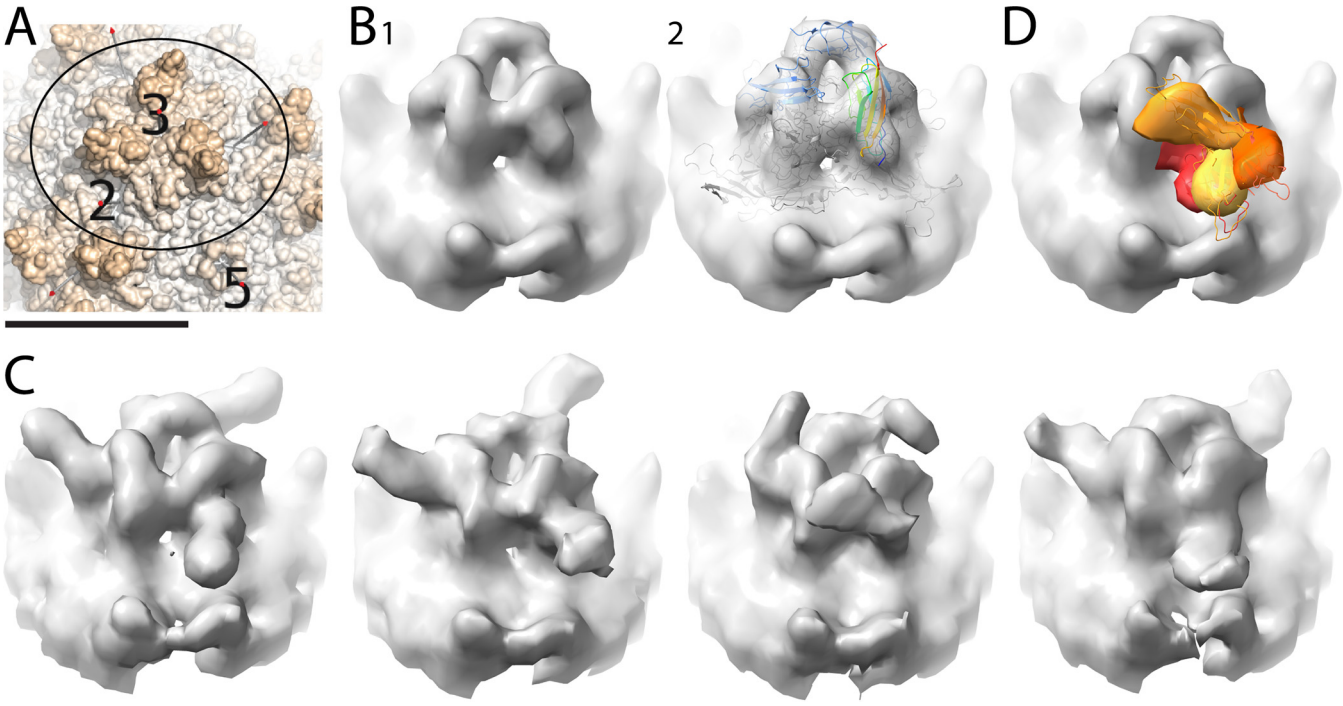

FIG 2 Classification of AAV2/PKD1-2 trimer spikes. (A) Orientating panel showing a surface representation of AAV2 from the atomic structure, PDB ID 1LP3. Scale bar: 100 Å. The circle shows the area that was classified by subvolume averaging, and the axes of symmetry are labeled as 2, 3, and 5. (B) Global tomographic subvolume average of AAV2/PKD1-2 trimeric spike as solid surface (1) and translucently (2) overlaid on the atomic model of a subunit trimer from the prior single-particle reconstruction (PDB ID 6nz0). The viral proteins are shown in gray ribbons, and two of the PKD2 domains are colored blue, while one is rainbow colored by residue number from N (blue) to C (red). (C) Classification of the region outside the PKD2 domain revealed 4 distinct conformations for PKD1. (D) The 4 PKD1 domain conformations were segmented and overlaid, C and PKD1 atomic models (shades of orange) were fit to the 4 classes seen in panel C.

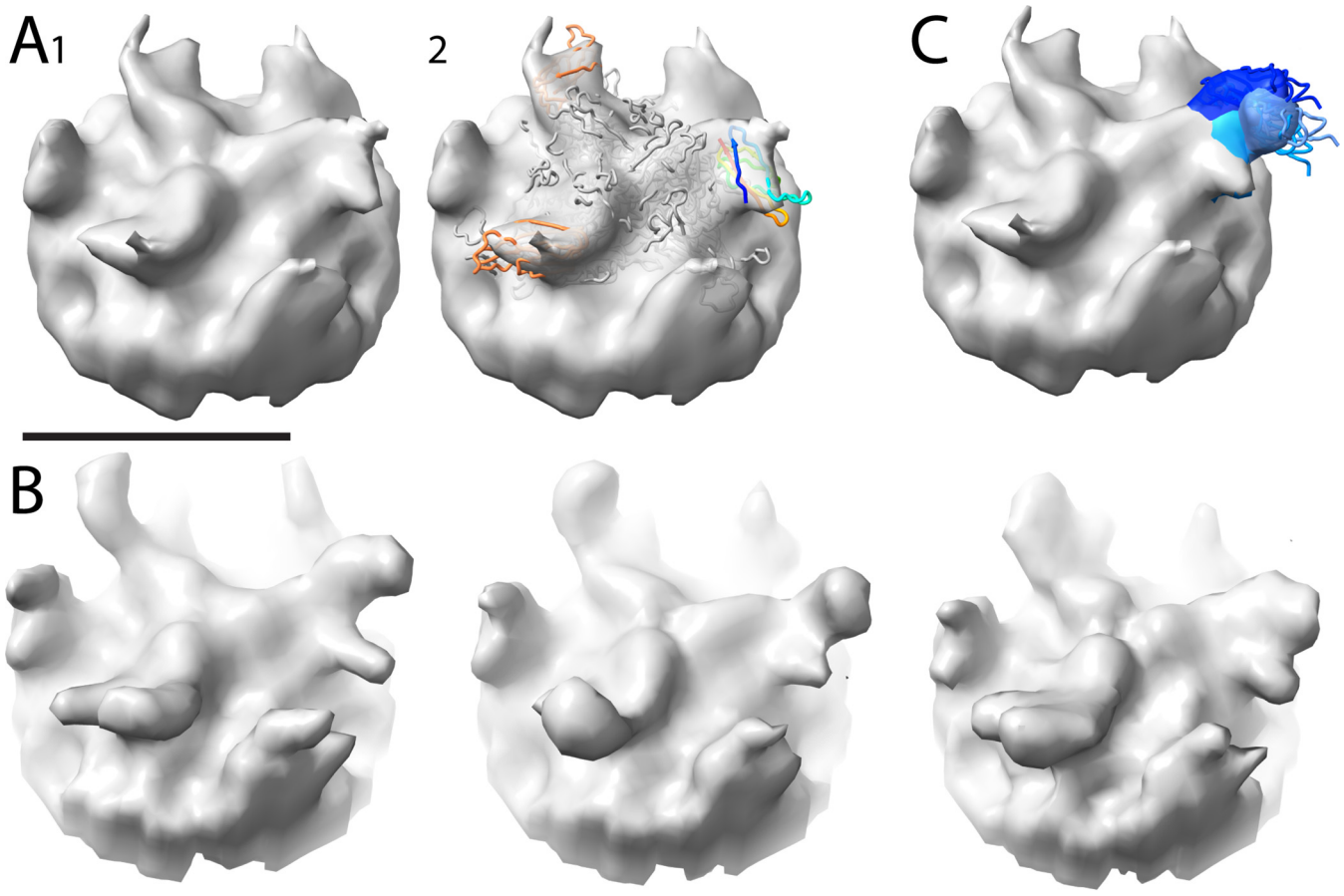

FIG 3 Classification of AAV5/PKD1-2 trimer spikes, oriented as in Fig. 2A. (A) Global tomographic subvolume average of the AAV5/PKD1-2 trimeric spike, alone (1) and overlaid (2) with the atomic structure from single-particle reconstruction (PDB ID 7kp3) in which the PKD1 domain of AAVR was seen. The viral proteins are shown in gray ribbons, and two of the PKD1 domains are colored orange, while one is rainbow colored by residue number from N (blue) to C (red). Scale bar: 100 Å. (B) Classification of the region outside the PKD1 domain revealed 3 distinct conformations for PKD2. (C) The 3 PKD2 domain conformations were segmented and are shown overlapping (shades of blue).

The structure of AAV5 bound with PKD1-2. Cryo-ET, subvolume averaging, and classification were performed on AAV5 in complex with the PKD1-2 domains similarly to those for AAV2/PKD1-2 (Fig. 1B and C). The global average of aligned subtomograms of AAV5 bound with PKD1-2 showed clear density for the PKD1 domain (Fig. 3A), in agreement with the previously published single-particle cryo-EM reconstructions of AAV5/PKD1-2 (59, 60). Assignment of the domain as PKD1 was by comparison of the cryo-ET reconstruction with high-resolution SPA, the latter showing a sequence distinctive to PKD1 and defining unambiguously the domain orientation with the N terminus near the viral 5-fold and the C terminus near the 2-fold. In the case of AAV5, cryo-ET and SPA were available not just for samples prepared in identical fashion, but from particles imaged from the same EM grid. Fitting the viral trimer spike model and PKD1 into the averaged tomographic map revealed features that were unaccounted for by the atomic models for the viral protein and PKD1 domain. These features were in positions extending away from the surface of the capsid that were plausible locations of PKD2. This differs from the AAV2/PDK1-2 tomography, where density for the "missing" domain (PKD1 for AAV2) was only revealed on classification and was not apparent in the global average. This indicates that the PKD2 domain is more constrained when AAVR is bound to AAV5 than PKD1 is when AAVR is bound to AAV2. The PKD2 domain becomes better defined upon classification, focusing on the area outside PKD1, which yielded three distinct classes (Fig. 3B). As with AAV2, the extra densities are the correct shape and size for a PKD domain. The length of the linker constrains the orientation with the N-terminal end of PKD2 to be close to the C terminus of PKD1,

TABLE 1 Consistency of the atomic models built into the AAV2-AAVR cryo-ET with cross-linking mass spectrometry (XL-MS)[a]

XL-MS expt					Distance measured from structure				
Virus	Virus residue	AAVR construct	Receptor residue	Location	Class 1 (Å)	Class 2 (Å)	Class 3 (Å)	Class 4 (Å)	Measured from:
AAVDJ	K558 (equivalent to AAV2 K556)	Full *ecto*-protein	K404	PKD2 (N-terminal)	13	13	13	13	PKD2 modeled into high-resolution EM of AAV2-PKD1/2
AAV2	K490	MBP-PKD1-5	K399	PKD1	37	30	26	15	PKD1 docked into cryo-ET
AAV2	T560	MBP-PKD1-5	399	PKD1	43	37	40	23	PKD1 docked into cryo-ET
AAV2	K556	MBP-PKD1-5	K338	PKD1	61	72	36	59	PKD1 docked into cryo-ET
AAV2	T450	MBP-PKD1-5	K399	PKD1	38	38	33	24	PKD1 docked into cryo-ET

[a]Under each class is listed the observed distance between the reactive groups on AAVR and AAV2.

and an atomic model of PKD2 fits well into the class maps (Fig. 3C). In each case, the first residue of the PKD2 model is within 19 Å of the C-terminal residue of PKD1, close enough to be bridged by the 5-residue domain linker. Variation in PKD2 orientation among classes of the AAV5 complex is modest (Fig. 3C), which is consistent with conformations that are more constrained than the AAVR conformations that we observed with AAV2.

Hybrid analysis: integration with cross-linking mass spectrometry (XL-MS). Meyer et al. reported mass spectroscopic identification of amino acids in AAV2 and AAV-DJ that were cross-linked with cyanurbiotindimercaptopropionylsuccinimide (CBDPS) that has a spacer length of 14 Å (44). Atomic models can be compared with these distance constraints, with the caveats that the tomographic classes are at ~20-Å resolution and without definition of side chains and that the distances are measured from cryo-ET samples that were not cross-linked and so do not reflect any remodeling of local structure on cross-linking (Table 1).

Given that, at best, nanometer-level consistency should be expected, class 4 provides a plausible explanation for 4 of 5 observed cross-links, 1 involving the tightly bound PKD2 and 3 involving the C terminus of PKD1, which is closest to PKD2 and the virus surface. A rationalization of K338 cross-linking (minimal 36 Å) is more tenuous, requiring remodeling of the lysine side chains. It seems more likely that tomography is sampling four of many possible PKD1 orientations and that any of the larger population could be captured in cross-linking. In other words, the cross-linking reflects a highly flexible receptor with many domain orientations, of which a subset, perhaps the most stable, are sampled in the tomographic classes.

Comparison of the AAV2 and AAV5 complexes with PKD1-2. When AAV2 is bound by PKD1-2, the PKD2 domain has the highest affinity, but PKD1 has measurable impact, while for AAV5, PKD1 appears to be the only domain involved (55). These results, coming from binding and transduction analysis of domain-swap and deletion mutants, are supported and rationalized by the current tomography study. The tomography reveals other differences in the PKD1/PKD2 domain modes of binding to the two serotypes. For PKD1-2 bound to AAV2, no density is revealed for the PKD1 domain in the global average of aligned subtomograms, indicating a high level of heterogeneity of the PKD1 domain. For the AAV5 complex, weak density of the previously missing PKD2 domain is apparent in the global average of the aligned subtomograms. This difference between global averages indicates that PKD1 in AAV2 is more heterogeneous than PKD2 in AAV5. This is further confirmed by the classes for the extra PKD1/PKD2. As shown in Fig. 4, the extent of variability in AAV2 is much higher than that of AAV5 with a wider range of orientations. Furthermore, for PKD1-2 bound to AAV2, three out of four of the classes are in extended conformations with obtuse angles between the two PKD domains, and the fourth class has the two PKD domains folded back on each other. For PKD1-2 bound to AAV5, the two domains are always at an acute angle, folded back toward one another and contacting near the hinge in an antiparallel hairpin configuration. The extra PKD1/PKD2 also differs on the contact with viral proteins. Consistent with the accessory role of PKD1 in AAV2 cellular entry, one the four classes

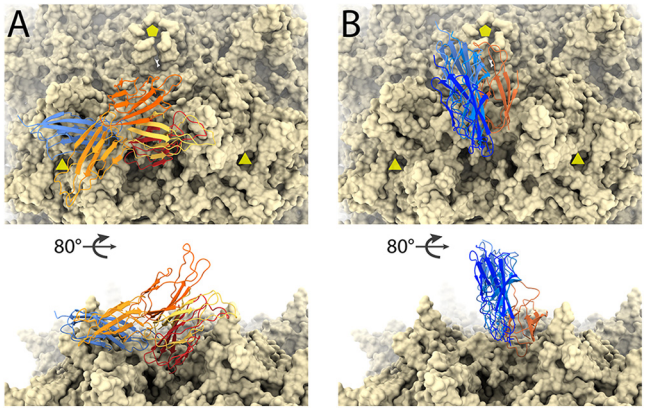

FIG 4 Conformational variability of AAV2/PDK1-2 and AAV5/PKD1-2 structures. The virus surfaces are viewed along a 2-fold axis (top) and tangentially (bottom), with 3-folds and 5-folds marked as triangles and pentagons, respectively. In each, the AAVR receptor is shown bound to one of the symmetry-equivalent regions on the virus surface. (A) AAV2, with the AAVR PKD2 domain (blue) tightly bound and the preceding PDK1 domain in several orientations (shades of orange) as determined from the subvolume tomography. (B) In its complex with AAV5, a tightly bound conformation of AAVR PKD1 (orange) is followed by PDK2 in three orientations (shades of blue) that are more constrained than PKD1 in panel A.

of PKD1 appears to have some contact with the viral protein, whereas for AAV5-bound PKD1, there appear to be no contacts with PKD2 (see below).

Contact between newly revealed extra PKD1/2 with VP protein. The details of interactions of PKD2 with AAV2 and PKD1 with AAV5 have been discussed in the respective high-resolution single-particle cryo-EM analyses (44, 60). Here, contact analyses are added for the newly revealed flexible PKD1 in its AAV2 complex and PKD2 in its AAV5 complex, but with a caveat that needs to be emphasized. The resolution of the tomographic classes (and therefore precision of atomic models) is low, ~20 Å. One must therefore be very cautious in interpreting whether 4.5-Å contact distance criteria are met for individual pairs of amino acids, and the analysis for each class is more appropriately an indicator of whether contacts are extensive or minimal. For the AAV2 class 4 model, only three residues of PKD1 (R353, V398, and K399) approach AAV2 closely (near E385 and D529 from different subunits). For the other three classes of AAV2 complex, there appear to be no contacts with PKD1 beyond the contacts already established at high resolution for the tightly bound PKD2 domain (44). These results are consistent with the cross-linking data in Table 1. For the classes of the AAV5 complex, no contacts are apparent between PKD2 and the viral protein.

Features on the inside surface of AAV2. In the global average of the aligned subtomograms of AAV2 bound with PKD1-2, there is extra density that projects in toward the center of the virus that is not accounted for by the known AAV atomic models (Fig. 5). Interestingly, the extra density was only observed in our AAV2/PKD1-2 structure (Fig. 5A) but not the AAV5/PKD1-2 structure (Fig. 5B). Model fitting shows that this protruding density is located in close proximity to residue 237 of the VP protein (Fig. 5C and E). Inspection of two previous AAV2/PKD1-2 single-particle cryo-EM reconstructions (which are otherwise similar) reveals differences in the maps in this same region (Fig. 5C and D). The Zhang et al. structure at 2.4 Å (EMD9672) is similar to most prior AAV structures, interpretable from residue 219, with β-strand A running antiparallel to βB before a hairpin turn at Gly_{236}-Asp_{237}, which connects the two strands (58). This map for βA is slightly weaker than those for βB and other strands, but only slightly, with only a slight hint of disorder (Fig. 5D). In contrast, in the 2.4-Å structure of Meyer et al. (EMD0553), there is little in the map to suggest the presence of βA (Fig. 5C) (44).

Relevant to these observations, it has been established that the AAV capsid is assembled from VP1, 2, and 3 in a roughly 1:1:10 ratio, sharing much of their sequence and structure but differing in their N-terminal extensions (see introductory section) (62). However, we have observed that the relative expression of VP1, 2, and 3 is variable in preparations for four different AAV serotypes. VP3 is always observed, but the

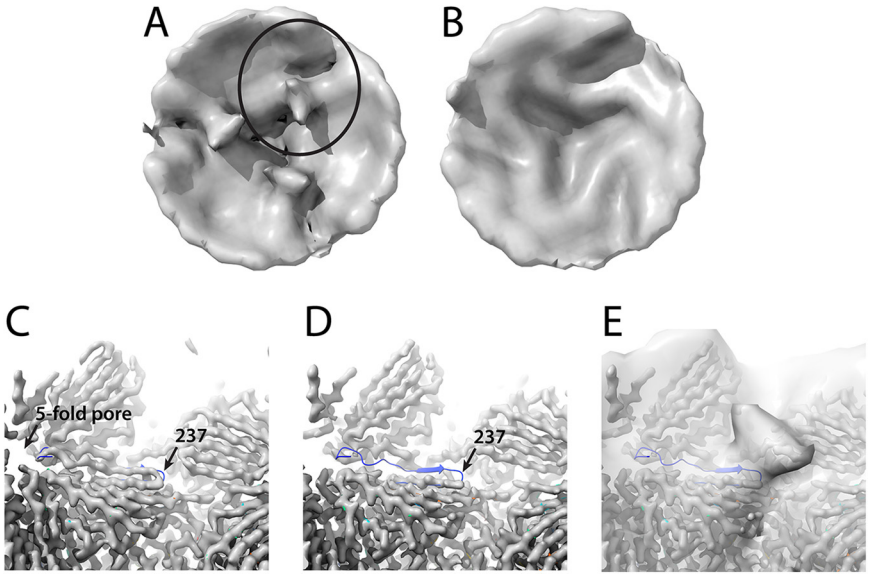

FIG 5 Unmodeled density inside the AAV2 capsid. (A) The inner surface of a subvolume average of the AAV2/PKD1-2 complex surrounding a 3-fold, viewed outwards. The circled region is highlighted in panels C to E. (B) The corresponding subvolume of the AAV5/PKD1-2 complex. (C) The 2.8-Å single-particle reconstruction (EMD9672) and model (PDB ID 6ihb) of an AAV2-AAVR (PKD1-5) complex (58). The map clearly shows βA proceeding left-to-right to the hairpin turn at residue 237 as has been seen in AAV2 crystal structures (1). (D) In the 2.4-Å single-particle reconstruction of an AAV2-AAVR complex (PKD1-2; EMD0553) (44), this configuration of βA is not apparent, as illustrated by overlaying the (mismatched) model PDB ID 6ihb (24, 58). (E) Overlaid on panel D is the newly observed AAV2/PKD1-2 subvolume average tomographic density, suggesting that the unmodeled density extending inward from the surface is a partially ordered configuration of upstream residues. It would be an alternate to the βA strand, connecting to βB at residue 237.

amounts of VP1 and VP2 vary. In addition, it has been shown that the relative incorporation of VP1 and VP2 during capsid assembly is stochastic, and the relative amounts vary considerably from capsid to capsid (63). It is possible that the proportions of VP1, 2, and 3 in individual particles, with their different N termini, is one of the factors affecting the fractions of subunits seen with ordered βA or disordered inward-pointing protrusions in different 3D reconstructions.

Disordered features in the same general area were first seen as "fuzzy globules" in a 2001 nanometer-resolution SPA reconstruction of empty AAV2 virus-like particles (64). Assignment as parts of VP1 and/or VP2 was supported by structures lacking fuzzy globules either for mutants in which VP1/2 were deleted or in capsids following heat treatment that was known to expose VP1u on the exterior of AAV (65). However, doubts emerged with the absence of the fuzzy globules in reconstructions of AAV1 vectors with various DNA contents, again by the same group (66). For the most part, these disordered features have not been noted in subsequent structures. However, they did resurface in the cryo-EM reconstruction of the AAV2 R432A mutant, in which map was missing at 3.7-Å resolution for βA. When this map was viewed at 5-Å resolution, there was a feature extending from the 236-7 hairpin turn that was interpreted as four residues extending toward the general area of the fuzzy globules (67). The map at 5-Å resolution had not been deposited, so for comparison to our cryo-ET, the 3.7-Å map (EMD8100) was low-pass-filtered to 11-Å resolution. At this lower resolution, we see not just 4 amino acids heading from the βA-βB turn toward the 2-fold (67), but additionally, we see the larger unmodeled feature that is also present in the AAV2/PKD12 tomography. Furthermore, this feature is the same as the fuzzy globules seen earlier in AAV2 VLPs at 11-Å resolution (64). For both our PKD1-2 complex and the R432A packaging mutant, higher-resolution cryo-EM SPA indicates that the presence of the unmodeled feature is accompanied by loss of βA; i.e., it is an alternative configuration for the N-terminal residues. In the prior 3.8-Å reconstruction of the wild-type AAV2, βA was clear, and we now

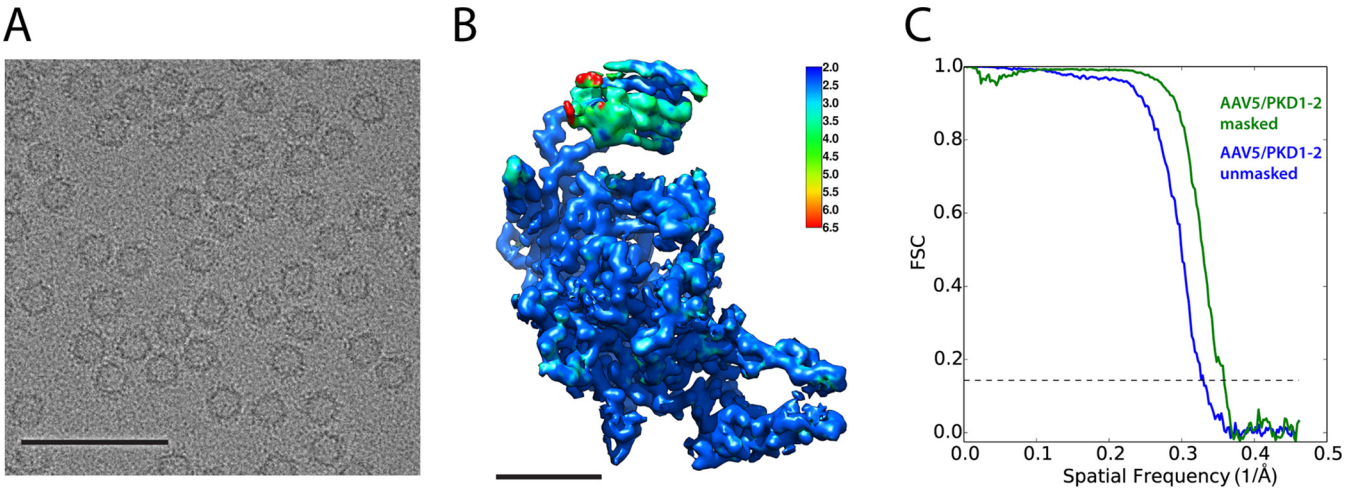

FIG 6 Single-particle analysis of AAV5/PKD1-2. (A) Representative micrograph of the single-particle AAV5/PKD1-2 data. Scale bar: 1,000 Å. (B) Protomer of the AAV5 PKD1-2 complex extracted from the SPA reconstruction colored by local resolution as estimated by ResMap. Scale bar: 25 Å. (C) Fourier shell correlation (FSC) curves for the AAV5/PKD1-2 SPA reconstruction. Curves are shown for the unmasked half maps (blue) and masked (green) with an estimated resolution of 2.8 Å.

add that when low-pass filtered to 11 Å, EMD8099 shows no indication of the partially ordered alternative configuration seen in the R432A mutant (67).

Whereas we see the partially ordered alternative configuration in our complex of AAV2 with the PKD1-2 fragment of AAVR, the nominally similar complex of AAV2 with a PKD1-5 fragment has density for βA that is only marginally weaker than that for βB (EMD 9671) (58), so the more usual configuration of βA predominates. Thus, there is not a simple and deterministic receptor-triggered conformational switch. Indeed weak, but recognizable density for Trp_{234} in the 2.4-Å structure of the PKD1-2 complex (44) indicates an equilibrium (favoring the alternative configuration) that might reflect incomplete PKD1-2 binding or an intrinsic and perhaps dynamic finely balanced equilibrium. The latter is consistent with a history of VLP structures where the alternative configuration is seen occasionally (64), but mostly not. So, we have nominally similar VLP and receptor-complex structures exhibiting different equilibrium states. Absent an understanding of how a finely balanced equilibrium is influenced, one should be cautious about attributing mechanistic significance, whether for packaging mutants or in receptor-binding. In summary, comparison of the Meyer et al. and Zhang et al. (44, 58) reconstructions reveals a putative conformational equilibrium in which N-terminally of Gly_{236}, the structures diverge. Tracing toward the N terminus, the chain either heads down βA toward the 5-fold pore or to disordered structures near the 2-fold pore (Fig. 5C to E). The volume of the partially ordered segment of the map is ~12,000 $Å^3$, corresponding to 89 typically sized amino acids. Thus, the disordered region, which is centered on a 2-fold axis, could contain two copies of the N-terminal 35 residues of VP3 (before βB). Alternatively, each could contain a single copy of either VP2 (65 + 35 = 100 residues before the βA-βB turn) or part of VP1 (202 + 35 = 237 residues), noting that there are a total of ~12 copies combined of VP1 and VP2, but 30 × 2-fold axes. For conformers headed toward the 5-fold, up to one in five would have access to the exterior through the pore, while others might be part of the "basket-like" disordered structure seen surrounding the 5-fold on the inner surfaces of some, but not all, AAVs (68).

Comparison of single-particle reconstructions for AAV5 bound with PKD1-2. A single-particle analysis of the AAV5/PKD1-2 was also performed using particles from the same grid that was subjected to tomographic analysis (Fig. 6A and B). The reconstruction, at 2.8- Å resolution (Fig. 6C), agrees well with the previous 2.5-Å map (60), similarly resolving PKD1, but not PKD2. The new SPA and tomography data were collected from the same sample grid, so detection of PKD2 is a result of the technique, not the sample. While the high-resolution SPA structures of PKD1 are mostly very similar, there are differences in the N-terminal residues (Fig. 7A, arrows). This is at the same region where the two prior

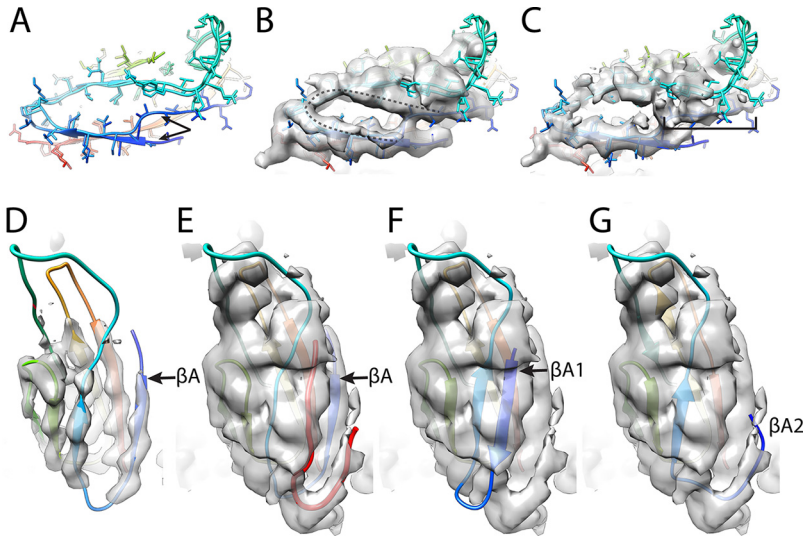

FIG 7 PKD1 structure from SPA of AAV5 complexes. (A) Comparison of the Silveria et al. (60) and Zhang et al. (58) models of AAVR PKD1 as bound to AAV5. Arrows indicate the discrepancies between the N-terminal residues of the two models. (B) Models in panel A compared to the cryo-EM SPA map determined here. The dashed line indicates unmodelled features at low contour. (C) Models in panel A compared to the Zhang et al. map, the bracket indicating part of the dashed region modeled by Zhang et al., but not Silveria et al. (D) At high contour, the current reconstruction supports the Silveria model as at least a dominant conformer. The first strand is labeled βA. (E) Low contouring of the new map shows two putative β strands (red) continuing the β sheets on the top and bottom of the PKD1 structure. In this model βA is in the same position as in (D). This panel shows The most probable interpretation, but panels F and G present an alternative. (F) Homologs in the PDB have structures similar to both the Silveria and Zhang models and, less frequently, to PDB ID 6aem. Overlay of this structure at 1.3-Å resolution shows how the longer of the unmodeled regions could be an alternative conformer of βA, labeled βA1. (G) The shorter unmodeled region would have to be a 3rd conformer of βA, labeled βA2.

single-particle analyses had been modeled differently (59, 60) in weak regions of their respective maps. The first β-strand of PKD1 is segmented, with βA followed by a break for a slight change of direction and then βA′. The structures diverge upstream of the break near residue 315 (Fig. 7A, arrows). The N-terminal residue of the Silveria et al. construct (60), Val_{311}, is not seen, but the rest of the structure is homologous to PKD PDB ID 2yrl. The Zhang et al. (58) construct starts upstream, so is longer at the N terminus. It is modeled more like homolog PDB ID 2y72, with βA occupying part of a different map feature and connecting to βA′ with a non-β linker (residues 315 to 313). This configuration allows the authors to predict a contact near the 5-fold (AAVR Val_{305} and AAV5 Ser_{319}), but the map at the N terminus is weak and does not offer experimental support. At low contour in an unsharpened map, the new 2.8-Å SPA reveals an additional extended chain/β-strand approximately parallel to βA′. The observation prompted retrospective examination of the prior 2.5- Å reconstruction (60). The same, hitherto unrecognized, feature looks very similar at a low-contour level in the unsharpened map. Furthermore, the map deposited by Zhang et al. was reexamined, and similar features were found (Fig. 7C), only some of which are accounted for by their different βA configuration. Given a lack of side chains, the identity of the extra features cannot be determined unambiguously. They are best described as a β-hairpin "U" (Fig. 7B). Hypothetically, one could account for the two arms of the U separately with two additional configurations for βA-βA′ of AAVR PKD1. The melting of β-sheet hydrogen bonds to spring βA′ loose seems implausible, and there is not the diminution of density for βA′ expected if βA′ had alternate conformers.

Using high contour levels for the new map, we see that the predominant configuration of PKD1 is as modeled by Silveria et al. (60) (Fig. 7). It is most likely that the new features belong to a single peptide distinct from that previously modeled. In the discussion, two possibilities will be considered: either that the unaccounted density is part of a second

AAVR subunit or that it is a fragment of hitherto unseen N-terminal regions of an AAV capsid protein.

DISCUSSION

Cryo-ET has allowed a subvolume classification that revealed the locations of receptor domains which were missing from the previous single-particle analyses (SPA), even with attempted SPA subvolume classification (44, 69). For AAV5, PKD1 had been visualized by SPA (59, 60), but there had been no sign of PKD2. The cryo-ET showed PKD2 doubled back over the top of PKD1 in two orientations differing by ~30°, likely neither of sufficient occupancy and order to be resolved by cryo-EM SPA. Such disorder and heterogeneity are consistent with PKD2 having few interactions, limited contacts with PKD1 near the interdomain hinge, and no contacts with AAV5 beyond the PKD1-2 domain linker for either of the classes. The distal locations of PKD2, revealed by cryo-ET, are also consistent with analysis of domain-deletion and chimeric domain-swapped mutants, which indicated that PKD1, but not PKD2, has significant impact upon AAV5 cellular transduction (55).

AAV2 presented more of an enigma, because the same mutational analysis found that PKD2 was most important for AAV2 entry, but PKD1 also enhanced transduction, though to lesser extent. PKD2, the more critical for transduction, had previously been resolved by SPA (44, 58), but the "accessory" PKD1 had not. Prior to the SPA structures of AAV5-AAVR complexes (59, 60), we hypothesized that the unseen PKD1 might be interacting loosely with AAV2 at a site corresponding to the (yet to be determined) AAV5/PKD1 interface. The cryo-ET shows that none of PKD1 locations of any of the four classes in the AAV2 complex bear any resemblance to PKD1 as bound by AAV5.

However, one of the four PKD1 classes has some direct contact with AAV2 proteins. This is consistent with PKD1 playing an accessory role not strictly required for, but enhancing, cellular transduction. Note, however, that only one of the four AAV2 classes appears to make contact, and the contact is not extensive. Thus, it is not surprising that there can be the observed wide-ranging heterogeneity in domain orientation, the four classes spanning a 120° rotation about the interdomain hinge. While we would expect the more populated orientations to rise to the top of classification, there might well be diversity beyond the four discretely classed orientations (as indicated by the XL-MS), and it is not surprising that PKD1 was not detectable by SPA. Clearly the level of interactions between PKD1 and either AAV2 or PKD2 are insufficient to restrict conformational heterogeneity, so one wonders whether the interactions with AAV2 can be strong enough to have a measurable direct impact upon transduction through avidity. It seems more likely that either PKD1 increases the availability or stability of AAVR in a state compatible with the binding of AAV2 to PKD1 or that there is a different step in AAV entry in which PKD1 has a role.

Completely unanticipated was the unmodeled density on the inside surface of the AAV2/PKD1-2 complex (but not the AAV5/PKD12 complex). It correlates inversely with the strength of βA density, density for βA being much weaker when the unmodeled features are seen. Thus, it appears that we are observing an equilibrium between two states, one with an ordered βA extending from the 5-fold region and the other with a partially ordered N-terminal region coming from the inner surface protrusion, skipping βA, and joining the jellyroll fold capsid protein at the βA-βB hairpin turn. The volume of the inner protrusion is commensurate with that expected of the N-terminal 35 residues of two VP3s meeting at a 2-fold axis, although one cannot rule out partial occupancy by VP2 or VP1. Whether and how this equilibrium in N terminus location is influenced by receptor-binding far away on the outside surface are unknown.

Another surprise was the previously unseen fragments of β-strand structure adjacent to PKD1 in its complex with AAV5. They lacked distinctive features to identify by sequence. Nevertheless, there are a limited number of plausible possibilities. The N-terminal regions of the capsid proteins have never been seen at high resolution. While in this study partially ordered structures were seen on the interior surface of AAV2, crystal structures of some

AAVs and autonomous parvoviruses have indicated that a fraction of N termini (of at least VP3) might be external: partially ordered density running down the 5-fold pore from the outside is interpreted as the connection to the start of the β-barrel on the inside surface (70–73). The absence of density on the 5-fold axis in the AAV5 single-particle analyses lessens the likelihood that the unaccounted features are previously unresolved N-terminal parts of the viral protein outside the capsid.

Alternatively, the extra peptides could come from unmodeled regions of AAVR. Dimers and higher oligomers are seen in preparations of PKD1-2 constructs (and MBP-PKD1-5 fusions) (44, 74). To date, AAVR dimers have not been observed bound to AAV5, but one cannot exclude the possibility that a small fraction of receptors in the complex are dimerized, with disorder that precludes EM observation of most of the second subunit.

This work is a testament to the value of combining multitechnique, multiscale approaches for flexible complexes, and in recognizing gaps in our understanding through exclusive reliance on high-resolution structure. A plan for multiple contingencies involved not only integration of different EM techniques, but also upstream redundancy in expression constructs, both of which were needed for a more robust and holistic understanding. It is noted that the first application of cryo-ET, to a complex of AAV2 with a PKD1-5 MBP fusion construct, led to a very low-resolution visualization that lacked domain definition or perception of conformational heterogeneity (44). It was only with a smaller construct, His_6-PKD1-2, that higher binding occupancy was achieved and conformational heterogeneity from domains 3 to 5 was eliminated, making it possible to classify the remaining heterogeneity and resolve distinct configurations for the two proximal domains. On the technical side, it is noted that fully automated classification of subvolume tomograms within a symmetrical particle was not yet possible. It is hoped that examples like this will inspire ongoing algorithm development, so that future applications will not be limited by the laboriousness of interactive classification.

MATERIALS AND METHODS

Virus and receptor preparation. Virus-like particles (VLPs) for both AAV2 and AAV5 were prepared as previously described (44, 60, 74). In brief, VLPs, which consist of the protein shells absent the viral DNA, were expressed in Sf9 cells using the Invitrogen Bac-to-Bac protocol and a pFastBacLIC cloning vector (Addgene no. 30111). VLPs then underwent three rounds of CsCl density gradient ultracentrifugation. An additional step of heparin affinity chromatography was performed on the AAV2 VLPs.

AAVR constructs were expressed using a pET-11a vector in *Escherichia coli* BL21(DE3) cells as described previously (44, 74). PKD1-2 was expressed with a 6× histidine tag and purified using Co^{2+} affinity chromatography.

Cryo-electron microscopy. Complexes of AAV2 with AAVR PKD1-2 were prepared on-grid as follows. A thin layer of carbon was deposited on a mica sheet using a Cressington carbon coater and was floated onto Quantifoil R1.2/1.3 grids. Then the carbon-coated Quantifoil grids were glow-discharged for 20 s with a Solarus 950 (Gatan). Then 4 μL of AAV2 VLP at a concentration of 0.2 mg/mL was applied on the grid and incubated for 1.5 min. The grid was gently blotted on the side with filter paper, and another 4 μL of PKD1-2 at a concentration of 1.5 mg/mL was applied and allowed to incubate for another 1.5 min followed by plunge-freeze using a Vitrobot Mark IV (FEI).

EM grids of AAV5 complexed with AAVR PKD1-2 were prepared in a similar way. VLP and receptor were first dialyzed into 25-mM HEPES and 150 mM NaCl, pH 7.4. EM samples were then prepared on glow-discharged ultrathin continuous carbon film supported by lacey carbon on copper grids (Ted Pella, Redding, CA; catalog [cat.] no. 01824). First, 2 μL of ~5.4 μM AAV5 VLP was added to the grid and given 2 min to adhere. Sample was then wicked, and 2 μL of 33 μM PKD1-2 was added. Grids were then plunge-frozen using an FEI Vitrobot Mark IV with a blot force of 4, time of 2 s, temperature of 25°C, and 100% humidity.

Cryo-ET tilt series were acquired on a Thermo Fisher Titan Krios (Hillsboro, OR) and recorded with Leginon software (75) on a Gatan K3 direct detector. A magnification of ×33,000 was used with a pixel size of 2.74 Å and a total dose of 100 $e^-/Å^2$ per tilt series. The tilt angle ranged from −60° to 60° with 2° steps. Exposure time at each tilt step was automatically adapted by the Leginon software according to the tilt angle. The number of frames at each tilt step was automatically set by Leginon according to the exposure time at each tilt step. The dose was fractionated across the frames at each step. Defocus values were set to 5 μm underfocus.

Single-particle were was also collected on the Titan Krios with the K3 camera using Leginon software. Magnification was ×81,000, and pixel size was 1.1 Å. The defocus range was set to −1.0 to −3.0 μm. The total dose was ~60 $e^-/Å^2$ per image with 50 frames for each micrograph. All frames of each micrograph were aligned using MotionCor2 (76).

For AAV5 bound with PKD1-2, both single-particle and tomography data were collected from the same cryo-EM grid.

Single-particle image processing. CTFFIND4 and GCTF were used to estimate contrast transfer function (CTF) parameters on all motion-corrected micrographs, and the best estimate was chosen using

resolution evaluation in Appion (77–79). Around 1,000 particles were picked using DoG (difference of Gaussian) Picker, and the rotational average of those particles was used as the template for picking using findEM in Appion (79, 80). A total of 26,091 particles were picked from 378 micrographs and extracted with a box size of 432 by 432 pixels in Appion. 2D classification and 3D classification were conducted to choose good particles in Relion3-beta. The previous 2.5-Å resolution single-particle cryo-EM reconstruction of the AAV5/PKD1-2 complex (EMD22988) was low-pass-filtered to 60-Å resolution and used as the initial reference for 3D refinement by Relion3-beta (44, 81). A total of 15,052 particles were selected for gold standard auto-refinement. Icosahedral symmetry was applied for auto-refinement. After auto-refinement, CTF refinement, and beam tilt refinement, a final map of 2.8-Å resolution was achieved (Fig. 6).

Tomography image processing. Tilt series were aligned using Protomo software within Appion (79, 82, 83). Following that, the image stack for tilt series was imported into EMAN2/e2tomo. Alignment parameters from Protomo were imported into EMAN2/e2tomo with home-made scripts, and tomograms were directly calculated with imported parameters. Contrast transfer function (CTF) parameters were estimated on all micrographs inside E2tomo. For AAV2 bound with PKD1-2, 127 virions (7,620 asymmetric subunits) were manually picked and extracted with E2tomo and a box size of 288 by 288 pixels. Similarly, 85 virions (5,100 asymmetric subunits) were picked and extracted for AAV5 bound with PKD1-2. The same 2.4-Å AAV2/PKD12 reconstruction (EMD-0553) was low-pass-filtered, now to 50 Å, and used as the initial reference for alignment. For the complex of PKD1-2 with AAV5, extracted 3D subtomograms were aligned using the new 2.8-Å SPA reconstruction of the PKD12-AAV5 complex and low-pass-filtered to 50 Å. Subtilt refinement was then used to align the individual 2D particle images in each tilt and apply a per-particle-per-tilt CTF correction. AAV2/PDK1-2 was refined to 20 Å (Fig. 1C, blue), and AAV5/PKD1-2 was refined to 25 Å (Fig. 1C, green). The subtilt refined 3D particles were exported from EMAN2/e2tomo and then imported into the program I3 (84, 85). Icosahedral symmetry was applied to generate 60 copies of each particle such that each possible asymmetric unit was overlaid onto the same frame of reference. The particles were then translated and rotated to center on the 3-fold spike in a "spike-up" standard orientation. At this point, particles were reextracted with a box size of 90 by 90 pixels surrounding a single 3-fold spike, facilitating classification of asymmetric units. In order to improve the signal/noise ratio, the reextracted subtomograms, containing one trimer, were binned by 2. Then classification was conducted on one asymmetric unit to reveal the PKD1 domain (AAV2 bound with PKD1-2) or PKD2 domain (AAV5 bound with PKD1-2). Even though the resulting classes are based on a single asymmetric unit, the classes were reexpanded by 3-fold symmetry to better illustrate the context of the extra domains.

Model fitting. All the tomography maps for AAV2 and AAV5 bound with PKD1-2 were aligned to the same frame of reference. High-resolution SPA reconstructions EMD0553 and EMD9672 were aligned with the global (overall) average of the subtomograms of AAV2 bound with PKD1-2 (44, 58). Similarly, the map for the newly obtained 2.8-Å single-particle reconstruction for the AAV5/PKD1-2 complex was aligned to the global average subtomogram of the same complex. PDB ID 6ihb, with its well-ordered βA, was used for the atomic model of the AAV2 viral protein, while PKD2 was taken from the higher resolution PDB ID 6nz0 (44, 58). Likewise, the atomic model of VP protein and PKD1 from PDB ID 7kpn was docked as a single rigid-body trimer into the newly obtained 2.8-Å reconstruction for the AAV5/PKD1-2 complex and used to interpret the tomography maps. In summary, the tomographic reconstructions for AAV2 and AAV5 complexes were calculated in the same frame of reference, and then high-resolution SPA reconstructions and atomic models were overlaid, facilitating comparisons.

The previously unseen domains were modeled as follows. Atomic models for PKD1 and PKD2 were taken from PDB entries 7kpn and 6nz0, respectively, and were fitted, as rigid domains, into the newly revealed domain densities separately for the 4 AAV2 classes and the 3 AAV5 classes using Chimera (86).

Contact analysis. The VMD atomselect command was used to identify additional potential residue contacts contributed by the AAVR domains that had not previously been resolved in single-particle analysis (87). Distances between the newly revealed PKD1 (AAV2)/PKD2 (AAV5) and respective viral proteins were calculated. The atomselect command lists the residue numbers of all residues that have any atom approaching within 4.5 Å.

Data availability. Maps for the subvolume averages and single-particle reconstruction have been deposited in the EM data resource database: EMD-26172—global average of aligned subtomograms of AAV2 bound with PKD1-2; EMD-26173—first class of AAV2 bound with PKD1-2 revealed by classification of aligned subtomograms; EMD-26174—second class of AAV2 bound with PKD1-2; EMD-26175—third class of AAV2 bound with PKD1-2; EMD-26176—fourth class of AAV2 bound with PKD1-2; EMD-26177—SPR reconstruction of AAV5 bound with PKD1-2; EMD-26182—global average of aligned subtomograms of AAV5 bound with PKD1-2; EMD-26186—first class of AAV5 bound with PKD1-2 revealed by classification of aligned subtomograms; EMD-26187—second class of AAV5 bound with PKD1-2; EMD-26189—third class of AAV5 bound with PKD1-2.

ACKNOWLEDGMENTS

Funding for this research was provided by the National Institutes of Health (R35GM122564) (M.S.C.) and in part by R01GM108753 (S.M.S.).

Conceptualization and methodology: S.M.S. and M.S.C. Funding acquisition: M.S.C. Sample preparation: M.A.S. Data collection: G.H. Data analysis: G.H. and M.A.S. Writing, original draft preparation: G.H., S.M.S., and M.S.C. Writing, review and editing: all authors.

REFERENCES

1. Xie Q, Bu W, Bhatia S, Hare J, Somasundaram T, Azzi A, Chapman MS. 2002. The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. Proc Natl Acad Sci U S A 99:10405–10410. https://doi.org/10.1073/pnas.162250899.
2. Caspar DLD, Klug A. 1962. Physical principles in the construction of regular viruses. Cold Spring Harbor Symp Quant Biol 27:1–24. https://doi.org/10.1101/sqb.1962.027.001.005.
3. Atchison RW, Casto BC, Hammon WM. 1965. Adenovirus-associated defective virus particles. Science 149:754–756. https://doi.org/10.1126/science.149.3685.754.
4. Conway J, Zolotukhin S, Muzyczka N, Hayward G, Byrne B. 1997. Recombinant adeno-associated virus type 2 replication and packaging is entirely supported by a herpes simplex virus type 1 amplicon expressing Rep and Cap. J Virol 71:8780–8789. https://doi.org/10.1128/JVI.71.11.8780-8789.1997.
5. Meier AF, Fraefel C, Seyffert M. 2020. The interplay between adeno-associated virus and its helper viruses. Viruses 12:662. https://doi.org/10.3390/v12060662.
6. Schlehofer JR, Dupressoir T. 2000. Infectiology and pathology of human adeno-associated viruses, p 59–67. *In* Faisst S, Rommelaere J (ed), Parvoviruses: from molecular biology to pathology and therapeutic uses. Karger, Basel, Switzerland.
7. Pattison JR. 1990. Parvoviruses: medical and biological aspects, p 1765–1784. *In* Fields BN, Knipe DM (ed), Virology. Raven Press, New York, NY.
8. Carter PJ, Samulski RJ. 2000. Adeno-associated viral vectors as gene delivery vehicles. Int J Mol Med 6:17–27. https://doi.org/10.3892/ijmm.6.1.17.
9. Carter BJ, Burstein H, Peluso RW. 2008. Adeno-associated virus and AAV vectors for gene delivery, p 115–156. *In* Templeton NS (ed), Gene and cell therapy: therapeutic mechanisms and strategies. CRC Press, Boca Raton, FL.
10. Russell DW, Grompe M. 2015. Adeno-associated virus finds its disease. Nat Genet 47:1104–1105. https://doi.org/10.1038/ng.3407.
11. Nault JC, Datta S, Imbeaud S, Franconi A, Mallet M, Couchy G, Letouze E, Pilati C, Verret B, Blanc JF, Balabaud C, Calderaro J, Laurent A, Letexier M, Bioulac-Sage P, Calvo F, Zucman-Rossi J. 2015. Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. Nat Genet 47:1187–1193. https://doi.org/10.1038/ng.3389.
12. La Bella T, Imbeaud S, Peneau C, Mami I, Datta S, Bayard Q, Caruso S, Hirsch TZ, Calderaro J, Morcrette G, Guettier C, Paradis V, Amaddeo G, Laurent A, Possenti L, Chiche L, Bioulac-Sage P, Blanc J-F, Letouze E, Nault J-C, Zucman-Rossi J. 2020. Adeno-associated virus in the liver: natural history and consequences in tumour development. Gut 69:737–747. https://doi.org/10.1136/gutjnl-2019-318281.
13. Donsante A, Miller DG, Li Y, Vogler C, Brunt EM, Russell DW, Sands MS. 2007. AAV vector integration sites in mouse hepatocellular carcinoma. Science 317:477. https://doi.org/10.1126/science.1142658.
14. Buning H, Schmidt M. 2015. Adeno-associated vector toxicity-to be or not to be? Mol Ther 23:1673–1675. https://doi.org/10.1038/mt.2015.182.
15. Nault J-C, Mami I, La Bella T, Datta S, Imbeaud S, Franconi A, Mallet M, Couchy G, Letouzé E, Pilati C, Verret B, Blanc J-F, Balabaud C, Calderaro J, Laurent A, Letexier M, Bioulac-Sage P, Calvo F, Zucman-Rossi J. 2016. Wild-type AAV insertions in hepatocellular carcinoma do not inform debate over genotoxicity risk of vectorized AAV. Mol Ther 24:660–661. https://doi.org/10.1038/mt.2016.47.
16. Berns KI, Byrne BJ, Flotte TR, Gao G, Hauswirth WW, Herzog RW, Muzyczka N, VandenDriessche T, Xiao X, Zolotukhin S, Srivastava A. 2015. Adeno-associated virus type 2 and hepatocellular carcinoma? Hum Gene Ther 26:779–781. https://doi.org/10.1089/hum.2015.29014.kib.
17. Dave UP, Cornetta K. 2021. AAV joins the rank of genotoxic vectors. Mol Ther 29:418–419. https://doi.org/10.1016/j.ymthe.2021.01.007.
18. de Jong YP, Herzog RW. 2021. Liver gene therapy and hepatocellular carcinoma: a complex web. Mol Ther 29:1353–1354. https://doi.org/10.1016/j.ymthe.2021.03.009.
19. Dalwadi DA, Torrens L, Abril-Fornaguera J, Pinyol R, Willoughby C, Posey J, Llovet JM, Lanciault C, Russell DW, Grompe M, Naugler WE. 2021. Liver injury increases the incidence of HCC following AAV gene therapy in mice. Mol Ther 29:680–690. https://doi.org/10.1016/j.ymthe.2020.10.018.
20. Maguire AM, Bennett J, Aleman EM, Leroy BP, Aleman TS. 2021. Clinical perspective: treating RPE65-associated retinal dystrophy. Mol Ther 29:442–463. https://doi.org/10.1016/j.ymthe.2020.11.029.
21. Mendell JR, Al-Zaidy S, Shell R, Arnold WD, Rodino-Klapac LR, Prior TW, Lowes L, Alfano L, Berry K, Church K, Kissel JT, Nagendran S, L'Italien J, Sproule DM, Wells C, Cardenas JA, Heitzer MD, Kaspar A, Corcoran S, Braun L, Likhite S, Miranda C, Meyer K, Foust KD, Burghes AHM, Kaspar BK. 2017. Single-dose gene-replacement therapy for spinal muscular atrophy. N Engl J Med 377:1713–1722. https://doi.org/10.1056/NEJMoa1706198.
22. Mendell JR, Al-Zaidy SA, Rodino-Klapac LR, Goodspeed K, Gray SJ, Kay CN, Boye SL, Boye SE, George LA, Salabarria S, Corti M, Byrne BJ, Tremblay JP. 2021. Current clinical applications of in vivo gene therapy with AAVs. Mol Ther 29:464–488. https://doi.org/10.1016/j.ymthe.2020.12.007.
23. Wilson JM, Flotte TR. 2020. Moving forward after two deaths in a gene therapy trial of myotubular myopathy. Hum Gene Ther 31:695–696. https://doi.org/10.1089/hum.2020.182.
24. Philippidis A. 2020. After third death, Audentes' AT132 remains on clinical hold. Hum Gene Ther 31:908–910. https://doi.org/10.1089/hum.2020.29133.bfs.
25. Paulk N. 2020. Gene therapy: it is time to talk about high-dose AAV: the deaths of two children with X-linked myotubular myopathy in the ASPIRO trial prompts a reexamination of vector safety. Genet Eng Biotechnol News 40:14–16. https://doi.org/10.1089/gen.40.09.04.
26. Srivastava A, Lusby EW, Berns KI. 1983. Nucleotide sequence and organization of the adeno-associated virus 2 genome. J Virol 45:555–564. https://doi.org/10.1128/JVI.45.2.555-564.1983.
27. Berns KI. 1996. *Parvoviridae*: the viruses and their replication, p 1017–1041. *In* Fields BN, Knipe DM, Howley PM (ed), Virology, 3rd ed Raven, Philadelphia, PA.
28. Hoque M, Ishizu K, Matsumoto A, Han SI, Arisaka F, Takayama M, Suzuki K, Kato K, Kanda T, Watanabe H, Handa H. 1999. Nuclear transport of the major capsid protein is essential for adeno- associated virus capsid formation. J Virol 73:7912–7915. https://doi.org/10.1128/JVI.73.9.7912-7915.1999.
29. Popa-Wagner R, Sonntag F, Schmidt K, King J, Kleinschmidt JA. 2012. Nuclear translocation of adeno-associated virus type 2 capsid proteins for virion assembly. J Gen Virol 93:1887–1898. https://doi.org/10.1099/vir.0.043232-0.
30. Farr GA, Zhang LG, Tattersall P. 2005. Parvoviral virions deploy a capsid-tethered lipolytic enzyme to breach the endosomal membrane during cell entry. Proc Natl Acad Sci U S A 102:17148–17153. https://doi.org/10.1073/pnas.0508477102.
31. Farr GA, Cotmore SF, Tattersall P. 2006. VP2 cleavage and the leucine ring at the base of the fivefold cylinder control pH-dependent externalization of both the VP1 N terminus and the genome of minute virus of mice. J Virol 80:161–171. https://doi.org/10.1128/JVI.80.1.161-171.2006.
32. Bleker S, Sonntag F, Kleinschmidt JA. 2005. Mutational analysis of narrow pores at the fivefold symmetry axes of adeno-associated virus type 2 capsids reveals a dual role in genome packaging and activation of phospholipase A2 activity. J Virol 79:2528–2540. https://doi.org/10.1128/JVI.79.4.2528-2540.2005.
33. Bleker S, Pawlita M, Kleinschmidt JA. 2006. Impact of capsid conformation and rep-capsid interactions on adeno-associated virus type 2 genome packaging. J Virol 80:810–820. https://doi.org/10.1128/JVI.80.2.810-820.2006.
34. Plevka P, Hafenstein S, Li L, D'Abrgamo A, Jr, Cotmore SF, Rossmann MG, Tattersall P. 2011. Structure of a packaging-defective mutant of minute virus of mice indicates that the genome is packaged via a pore at a 5-fold axis. J Virol 85:4822–4827. https://doi.org/10.1128/JVI.02598-10.
35. Gao G, Vandenberghe LH, Alvira MR, Lu Y, Calcedo R, Zhou X, Wilson JM. 2004. Clades of adeno-associated viruses are widely disseminated in human tissues. J Virol 78:6381–6388. https://doi.org/10.1128/JVI.78.12.6381-6388.2004.
36. Zinn E, Vandenberghe LH. 2014. Adeno-associated virus: fit to serve. Curr Opin Virol 8:90–97. https://doi.org/10.1016/j.coviro.2014.07.008.
37. Harbison CE, Weichert WS, Gurda BL, Chiorini JA, Agbandje-McKenna M, Parrish CR. 2012. Examining the cross-reactivity and neutralization mechanisms of a panel of mAbs against adeno-associated virus serotypes 1 and 5. J Gen Virol 93:347–355. https://doi.org/10.1099/vir.0.035113-0.
38. Mietzsch M, Broecker F, Reinhardt A, Seeberger PH, Heilbronn R. 2014. Differential adeno-associated virus serotype-specific interaction patterns with synthetic heparins and other glycans. J Virol 88:2991–3003. https://doi.org/10.1128/JVI.03371-13.
39. Gao G, Lu Y, Calcedo R, Grant RL, Bell P, Wang L, Figueredo J, Lock M, Wilson JM. 2006. Biology of AAV serotype vectors in liver-directed gene transfer to nonhuman primates. Mol Ther 13:77–87. https://doi.org/10.1016/j.ymthe.2005.08.017.
40. Pillay S, Meyer NL, Puschnik AS, Davulcu O, Diep J, Ishikawa Y, Jae LT, Wosen JE, Nagamine CM, Chapman MS, Carette JE. 2016. An essential receptor for adeno-associated virus infection. Nature 530:108–112. https://doi.org/10.1038/nature16465.
41. Summerford C, Samulski RJ. 1998. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. J Virol 72:1438–1445. https://doi.org/10.1128/JVI.72.2.1438-1445.1998.

42. Kaludov N, Brown KE, Walters RW, Zabner J, Chiorini JA. 2001. Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. J Virol 75:6884–6893. https://doi.org/10.1128/JVI.75.15.6884-6893.2001.
43. Huang LY, Halder S, Agbandje-McKenna M. 2014. Parvovirus glycan interactions. Curr Opin Virol 7:108–118. https://doi.org/10.1016/j.coviro.2014.05.007.
44. Meyer NL, Hu G, Davulcu O, Xie Q, Noble AJ, Yoshioka C, Gingerich DS, Trzynka A, David L, Stagg SM, Chapman MS. 2019. Structure of the gene therapy vector, adeno-associated virus with its cell receptor, AAVR. Elife 8:e44707. https://doi.org/10.7554/eLife.44707.
45. Qing K, Mah C, Hansen J, Zhou S, Dwarki V, Srivastava A. 1999. Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. Nat Med 5:71–77. https://doi.org/10.1038/4758.
46. Goldman MJ, Wilson JM. 1995. Expression of alpha v beta 5 integrin is necessary for efficient adenovirus-mediated gene transfer in the human airway. J Virol 69:5951–5958. https://doi.org/10.1128/JVI.69.10.5951-5958.1995.
47. Summerford C, Bartlett JS, Samulski RJ. 1999. AlphaVbeta5 integrin: a co-receptor for adeno-associated virus type 2 infection. Nat Med 5:78–82. https://doi.org/10.1038/4768.
48. Asokan A, Hamra JB, Govindasamy L, Agbandje-McKenna M, Samulski RJ. 2006. Adeno-associated virus type 2 contains an integrin alpha5beta1 binding domain essential for viral cell entry. J Virol 80:8961–8969. https://doi.org/10.1128/JVI.00843-06.
49. Kashiwakura Y, Tamayose K, Iwabuchi K, Hirai Y, Shimada T, Matsumoto K, Nakamura T, Watanabe M, Oshimi K, Daida H. 2005. Hepatocyte growth factor receptor is a coreceptor for adeno-associated virus type 2 infection. J Virol 79:609–614. https://doi.org/10.1128/JVI.79.1.609-614.2005.
50. Weller ML, Amornphimoltham P, Schmidt M, Wilson PA, Gutkind JS, Chiorini JA. 2010. Epidermal growth factor receptor is a co-receptor for adeno-associated virus serotype 6. Nat Med 16:662–664. https://doi.org/10.1038/nm.2145.
51. Di Pasquale G, Davidson BL, Stein CS, Martins I, Scudiero D, Monks A, Chiorini JA. 2003. Identification of PDGFR as a receptor for AAV-5 transduction. Nat Med 9:1306–1312. https://doi.org/10.1038/nm929.
52. Akache B, Grimm D, Pandey K, Yant SR, Xu H, Kay MA. 2006. The 37/67-kilodalton laminin receptor is a receptor for adeno-associated virus serotypes 8, 2, 3, and 9. J Virol 80:9831–9836. https://doi.org/10.1128/JVI.00878-06.
53. Meisen WH, Nejad ZB, Hardy M, Zhao H, Oliverio O, Wang S, Hale C, Ollmann MM, Collins PJ. 2020. Pooled screens identify GPR108 and TM9SF2 as host cell factors critical for AAV transduction. Mol Ther Methods Clin Dev 17:601–611. https://doi.org/10.1016/j.omtm.2020.03.012.
54. Dudek AM, Zabaleta N, Zinn E, Pillay S, Zengel J, Porter C, Franceschini JS, Estelien R, Carette JE, Zhou GL, Vandenberghe LH. 2020. GPR108 is a highly conserved AAV entry factor. Mol Ther 28:367–381. https://doi.org/10.1016/j.ymthe.2019.11.005.
55. Pillay S, Zou W, Cheng F, Puschnik AS, Meyer NL, Ganaie SS, Deng X, Wosen JE, Davulcu O, Yan Z, Engelhardt JF, Brown KE, Chapman MS, Qiu J, Carette JE. 2017. Adeno-associated virus (AAV) serotypes have distinctive interactions with domains of the cellular AAV receptor. J Virol 91:e00391-17. https://doi.org/10.1128/JVI.00391-17.
56. Bycroft M, Bateman A, Clarke J, Hamill SJ, Sandford R, Thomas RL, Chothia C. 1999. The structure of a PKD domain from polycystin-1: implications for polycystic kidney disease. EMBO J 18:297–305. https://doi.org/10.1093/emboj/18.2.297.
57. Poon MW, Tsang WH, Chan SO, Li HM, Ng HK, Waye MM. 2011. Dyslexia-associated kiaa0319-like protein interacts with axon guidance receptor nogo receptor 1. Cell Mol Neurobiol 31:27–35. https://doi.org/10.1007/s10571-010-9549-1.
58. Zhang R, Cao L, Cui M, Sun Z, Hu M, Zhang R, Stuart W, Zhao X, Yang Z, Li X, Sun Y, Li S, Ding W, Lou Z, Rao Z. 2019. Adeno-associated virus 2 bound to its cellular receptor AAVR. Nat Microbiol 4:675–682. https://doi.org/10.1038/s41564-018-0356-7.
59. Zhang R, Xu G, Cao L, Sun Z, He Y, Cui M, Sun Y, Li S, Li H, Qin L, Hu M, Yuan Z, Rao Z, Ding W, Rao Z, Lou Z. 2019. Divergent engagements between adeno-associated viruses with their cellular receptor AAVR. Nat Commun 10:3760. https://doi.org/10.1038/s41467-019-11668-x.
60. Silveria MA, Large EE, Zane GM, White TA, Chapman MS. 2020. The structure of an AAV5-AAVR complex at 2.5 A resolution: implications for cellular entry and immune neutralization of AAV gene therapy vectors. Viruses 12:1326. https://doi.org/10.3390/v12111326.
61. Hu G, Liu J, Roux KH, Taylor KA. 2017. Structure of simian immunodeficiency virus envelope spikes bound with CD4 and monoclonal antibody 36D5. J Virol 91:e00134-17. https://doi.org/10.1128/JVI.00134-17.
62. Muzyczka N, Berns KI. 2001. Parvoviridae: the viruses and their replication, p 2327–2360. *In* Fields BN, Knipe DM, Howley PM (ed), Virology, 4 ed Lippincott Williams & Wilkins, Philadelphia, PA.
63. Wörner TP, Bennett A, Habka S, Snijder J, Friese O, Powers T, Agbandje-McKenna M, Heck AJR. 2021. Adeno-associated virus capsid assembly is divergent and stochastic. Nat Commun 12:1642. https://doi.org/10.1038/s41467-021-21935-5.
64. Kronenberg S, Kleinschmidt JA, Bottcher B. 2001. Electron cryo-microscopy and image reconstruction of adeno-associated virus type 2 empty capsids. EMBO Rep 2:997–1002. https://doi.org/10.1093/embo-reports/kve234.
65. Kronenberg S, Bottcher B, von der Lieth CW, Bleker S, Kleinschmidt JA. 2005. A conformational change in the adeno-associated virus type 2 capsid leads to the exposure of hidden VP1 N termini. J Virol 79:5296–5303. https://doi.org/10.1128/JVI.79.9.5296-5303.2005.
66. Gerlach B, Kleinschmidt JA, Bottcher B. 2011. Conformational changes in adeno-associated virus type 1 induced by genome packaging. J Mol Biol 409:427–438. https://doi.org/10.1016/j.jmb.2011.03.062.
67. Drouin LM, Lins B, Janssen M, Bennett A, Chipman P, McKenna R, Chen W, Muzyczka N, Cardone G, Baker TS, Agbandje-McKenna M. 2016. Cryo-electron microscopy reconstruction and stability studies of the wild type and the R432A variant of adeno-associated virus type 2 reveal that capsid structural stability is a major factor in genome packaging. J Virol 90:8542–8551. https://doi.org/10.1128/JVI.00575-16.
68. Mietzsch M, Barnes C, Hull JA, Chipman P, Xie J, Bhattacharya N, Sousa D, McKenna R, Gao G, Agbandje-McKenna M. 2020. Comparative analysis of the capsid structures of AAVrh.10, AAVrh.39, and AAV8. J Virol 94:e01769-19. https://doi.org/10.1128/JVI.01769-19.
69. Ilca SL, Kotecha A, Sun X, Poranen MM, Stuart DI, Huiskonen JT. 2015. Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. Nat Commun 6:8843. https://doi.org/10.1038/ncomms9843.
70. Nam HJ, Lane MD, Padron E, Gurda B, McKenna R, Kohlbrenner E, Aslanidi G, Byrne B, Muzyczka N, Zolotukhin S, Agbandje-McKenna M. 2007. Structure of adeno-associated virus serotype 8, a gene therapy vector. J Virol 81:12260–12271. https://doi.org/10.1128/JVI.01304-07.
71. Xie Q, Chapman MS. 1996. Canine parvovirus capsid structure, analyzed at 2.9 A resolution. J Mol Biol 264:497–520. https://doi.org/10.1006/jmbi.1996.0657.
72. Agbandje-McKenna M, Llamas-Saiz AL, Wang F, Tattersall P, Rossmann MG. 1998. Functional implications of the structure of the murine parvovirus, minute virus of mice. Structure 6:1369–1381. https://doi.org/10.1016/s0969-2126(98)00137-3.
73. Kaufmann B, Chipman PR, Kostyuchenko VA, Modrow S, Rossmann MG. 2008. Visualization of the externalized VP2 N termini of infectious human parvovirus B19. J Virol 82:7306–7312. https://doi.org/10.1128/JVI.00512-08.
74. Meyer N, Davulcu O, Xie Q, Silveria M, Zane G, Large E, Chapman M. 2020. Expression and purification of adeno-associated virus virus-like particles in a baculovirus system and AAVR ectodomain constructs in E. coli. Bio Protoc 10:e3513. https://doi.org/10.21769/BioProtoc.3513.
75. Suloway C. et al. (2009). Fully automated, sequential tilt-series acquisition with Leginon. J Struct Biol 167:11–18.
76. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods 14:331–332. https://doi.org/10.1038/nmeth.4193.
77. Rohou A, Grigorieff N. 2015. CTFFIND4: fast and accurate defocus estimation from electron micrographs. J Struct Biol 192:216–221. https://doi.org/10.1016/j.jsb.2015.08.008.
78. Zhang K. 2016. Gctf: real-time CTF determination and correction. J Struct Biol 193:1–12. https://doi.org/10.1016/j.jsb.2015.11.003.
79. Lander GC, Stagg SM, Voss NR, Cheng A, Fellmann D, Pulokas J, Yoshioka C, Irving C, Mulder A, Lau PW, Lyumkis D, Potter CS, Carragher B. 2009. Appion: an integrated, database-driven pipeline to facilitate EM image processing. J Struct Biol 166:95–102. https://doi.org/10.1016/j.jsb.2009.01.002.
80. Voss NR, Yoshioka CK, Radermacher M, Potter CS, Carragher B. 2009. DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. J Struct Biol 166:205–213. https://doi.org/10.1016/j.jsb.2009.01.004.
81. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM

structure determination in RELION-3. Elife 7:e42166. https://doi.org/10.7554/eLife.42166.
82. Winkler H, Taylor KA. 2006. Accurate marker-free alignment with simultaneous geometry determination and reconstruction of tilt series in electron tomography. Ultramicroscopy 106:240–254.
83. Noble AJ, Stagg SM. 2015. Automated batch fiducial-less tilt-series alignment in Appion using Protomo. J Struct Biol 192:270–278.
84. Winkler H. 2007. 3D reconstruction and processing of volumetric data in cryo-electron tomography. J Struct Biol 157:126–137. https://doi.org/10.1016/j.jsb.2006.07.014.
85. Winkler H, Zhu P, Liu J, Ye F, Roux KH, Taylor KA. 2009. Tomographic subvolume alignment and subvolume classification applied to myosin V and SIV envelope spikes. J Struct Biol 165:64–77. https://doi.org/10.1016/j.jsb.2008.10.004.
86. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera: a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612. https://doi.org/10.1002/jcc.20084.
87. Humphrey W, Dalke A, Schulten K. 1996. VMD: Visual molecular dynamics. Journal of Molecular Graphics 14:33–38.