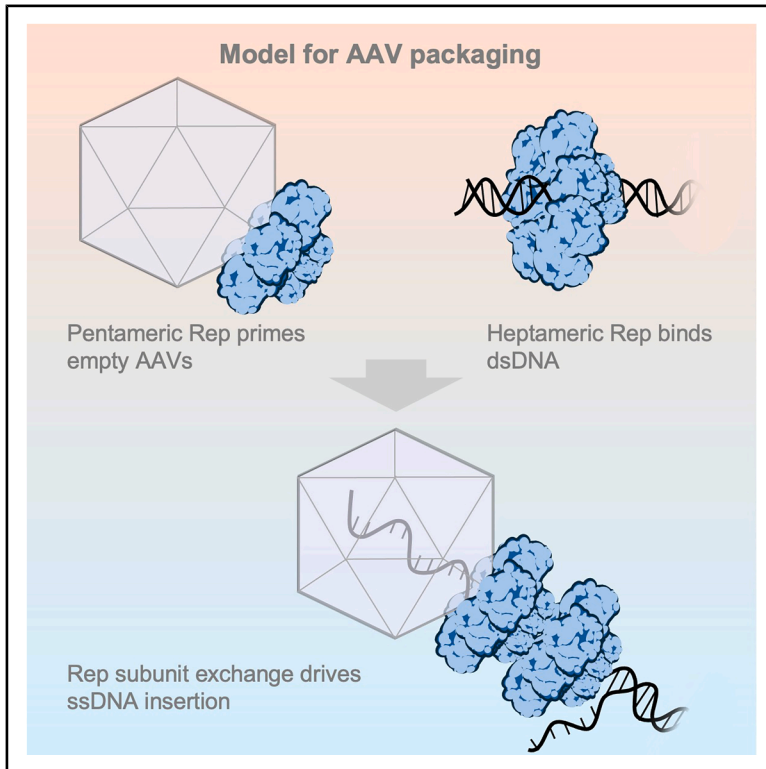


Structural basis for Rep-mediated adeno-associated virus packaging

Graphical abstract



Authors

Victoria Xu, Aiysha McInnes, Matthew Wake, Marta Acebrón-García-de-Eulate, Joseph D. Barritt, Doryen Bubeck, Sarah L. Rouse

Correspondence

joseph.barritt@astx.com (J.D.B.), d.bubeck@imperial.ac.uk (D.B.), s.rouse@imperial.ac.uk (S.L.R.)

In brief

Xu et al. report cryo-EM structures of the AAV packaging motor bound to the empty capsid and marker DNA. The motor-capsid complex reveals a hidden ATP-binding site as part of capsid structural changes pre-packaging, while the motor-DNA complexes show how oligomeric transitions drive DNA unwinding and insertion during genome packaging.

Highlights

- Cryo-EM structures of AAV2 Rep helicase bound to ITR-DNA and empty AAV8 capsid
- Rep-ITR complexes reveal DNA unwinding and translocation via conformational analysis
- Rep-capsid complex defines a binding interface likely conserved across parvoviruses
- Rep-capsid complex reveals an ATP binding site that primes the capsid for packaging



Article

Structural basis for Rep-mediated adeno-associated virus packaging

Victoria Xu,^{1,2} Aiysha McInnes,¹ Matthew Wake,² Marta Acebrón-García-de-Eulate,² Joseph D. Barritt,^{1,3,*} Doryen Bubeck,^{1,*} and Sarah L. Rouse^{1,4,*}

¹Department of Life Sciences, Imperial College London, London SW7 2AZ, UK

²AstraZeneca, Discovery Sciences, BioPharmaceuticals R&D Unit, Cambridge CB2 0AA, UK

³Present address: Astex Pharmaceuticals, 436 Cambridge Science Park, Milton Road, Cambridge CB4 0QA, UK

⁴Lead contact

*Correspondence: joseph.barritt@astx.com (J.D.B.), d.bubeck@imperial.ac.uk (D.B.), s.rouse@imperial.ac.uk (S.L.R.)

<https://doi.org/10.1016/j.celrep.2026.117044>

SUMMARY

Adeno-associated viruses (AAVs) are parvoviruses utilized as gene therapy vectors. However, the AAV packaging mechanism is unresolved at the molecular level, creating a bottleneck for vector manufacturing, safety, and efficacy. Here, cryo-EM structures of the Rep helicase packaging motor in complex with the packaging marker DNA (ITR) and the Rep-AAV8 capsid complex are presented. Rep-ITR complexes reveal dynamic oligomeric states on the DNA, elucidating the strand separation mechanism coupled to its ATPase cycle. We observe Rep preferentially bound to empty capsids, with a binding interface likely conserved across the virus family. This complex also unveils a cryptic capsid ATP-binding site which, alongside Rep binding, triggers structural rearrangements priming the capsid for packaging. Collectively, these findings advance the understanding of Rep-mediated packaging, with significant implications for parvovirus virology and viral vector design.

INTRODUCTION

Adeno-associated viruses (AAVs) use replication (Rep) motors to package DNA for gene therapy. AAVs are primate-infecting, non-pathogenic viruses, making them ideal viral vectors for healthcare applications.^{1–3} Indeed, there are 7 FDA/EMA-approved AAV-based gene therapy products and over 230 clinical trials worldwide.⁴ Advancements in vector engineering have improved efficacy and tissue specificity,^{5,6} while suspension cell lines offer improved yield.^{7,8} However, AAV packaging is an inefficient process that varies with DNA payload, host cell, and helper virus identity.^{9–13} Inefficient packaging combined with manufacturing challenges pose major barriers to the widespread application of AAV-based therapies. Consequently, recent studies have targeted Rep motors for engineering to enhance packaging,^{9,14–16} despite incomplete understanding of the process. Molecular understanding of DNA packaging is therefore crucial for our ability to increase safety, efficacy, and reduce the production cost for AAV-based gene therapies.

As members of the non-enveloped virus *Parvoviridae* family, AAVs package by inserting their ssDNA genome into empty capsids via molecular motors.¹⁷ The AAV genome is flanked by inverted-terminal-repeats (ITRs), which fold into DNA hairpins (Figure S1A). They serve as “packaging signals” for the DNA.^{18–21} Within the genome, open reading frames encode non-structural proteins^{22–24} and structural capsid proteins (VP1/2/3)²⁵ (Figure S1B). The non-structural Rep proteins are central to virus generation. There are four isoforms of Rep where large Reps (78/68) nick and unwind DNA during AAV genome

replication, while small Reps (52/40) package single-stranded DNA into empty capsids.^{17,26} They share a core helicase domain, comprising oligomerization and AAA+ ATPase subdomains (Figure S1C). Consistent with their role during replication, large Reps have an additional origin binding domain that connects to the helicase domain through an oligomerization linker.²⁷ Meanwhile, zinc-finger domains not involved in DNA binding can be found in both small and large Reps (Rep52/Rep78).²⁸

Reps are involved in transcription, replication, and packaging stages of the viral life cycle. AAV replication depends on helper virus proteins to regulate the cell cycle, while host machinery and Rep proteins coordinate DNA synthesis and processing.^{29,30} Large Reps with origin binding domains specifically bind the Rep binding element on the ITR, subsequently recruiting cellular replication complexes such as DNA polymerase δ and the MCM complex.^{30–32} Rep and host helicases coordinate unwinding during replication.³¹ Furthermore, large Reps facilitate DNA processing by nicking the terminal resolution site and covalently attaching to the new 5' end.^{33,34} These tasks are performed under different oligomeric states. Structural studies show the Rep68 origin binding domain forms trimers-of-dimers, while its helicase domain forms heptamers on double-stranded AAV integration site DNA.^{35,36} On single-stranded DNA, the origin binding domain are monomers or octamers, while helicase domains form heptamers or hexamers.^{35,37} Following replication, small Reps then act as motors that work from 3' to 5' to pump the ITR-DNA across the 5-fold capsid pore.^{17,38,39} Crystallographic structures of monomeric Rep40 bound to ADP and Mg^{2+} revealed that ATP hydrolysis requires oligomerization.^{40,41}



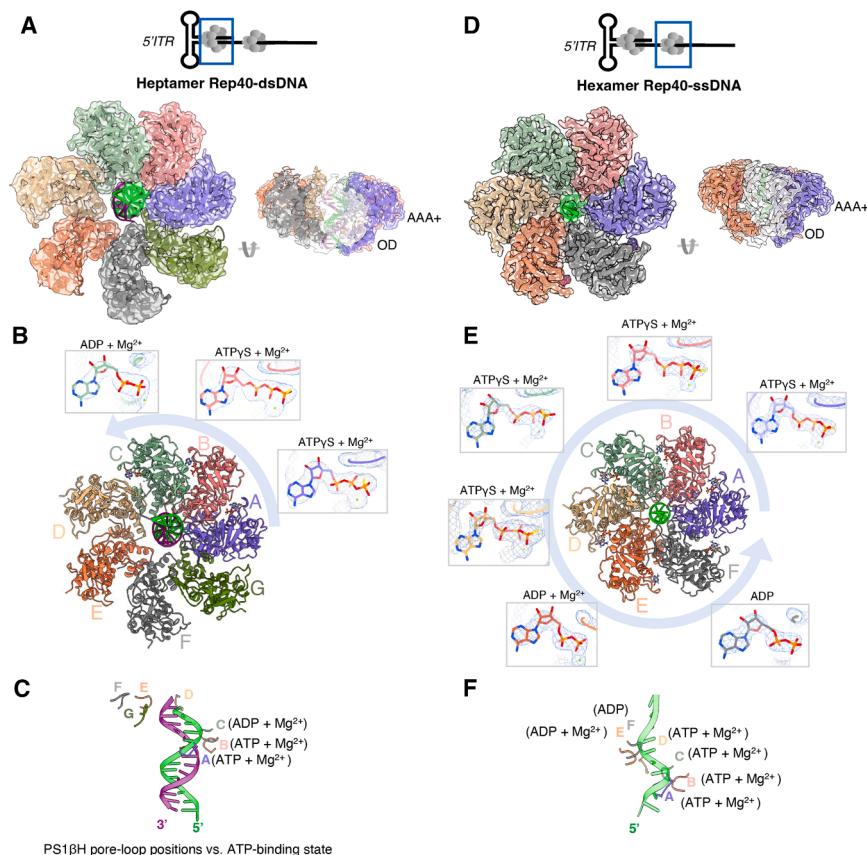


Figure 1. Cryo-EM structure of AAV2 Rep40 bound to 5' ITR

(A) Map-model overlay of heptameric Rep40 in complex with the dsDNA region of ITR. Two views rotated 90°. Three compacted subunits contact dsDNA (passive strand, purple; tracking strand, green). AAA+ and oligomerization domains (OD) of Rep40 are indicated.

(B) Three ATP-binding sites bridge the compacted subunits in the heptamer. Arrow indicates the ATPase cycle. Subunits are labeled and colored as in (A). Density for each ATP_γS is shown (inset).

(C) DNA-binding loop (pre-sensor-1-β-hairpin [PS1βH]) of the heptamer positioned along dsDNA. ATP-binding state for each heptamer subunit is labeled.

(D) Map-model overlay of hexameric Rep40 in complex with the ssDNA region of ITR.

(E) Six ATP-binding sites bridge the compacted subunits in the hexamer. Arrow indicates the ATPase cycle. Subunits are labeled and colored as in (D). Density for each ATP_γS is shown (inset).

(F) DNA-binding loop (PS1βH) of the hexamer positioned along ssDNA. ATP-binding state for each hexamer subunit is labeled.

See also [Figures S1–S3](#).

However, how oligomeric small Reps unwind and insert ITR-DNA during packaging remains unresolved. Furthermore, interaction studies show that the Rep helicase domain and capsid form a “pre-encapsidation complex” prior to DNA insertion.^{17,26} However, how small Reps bind the capsids and ITR-DNA for packaging remains unknown at the structural level.

In this study, we set out to understand how the small Rep (Rep40) engages both the ITR and the capsid to facilitate packaging. We use cryo-electron microscopy (cryo-EM) to determine the structure of Rep40 of AAV2, a serotype used ubiquitously in therapeutic AAV production. We present a series of Rep40 structures bound to the ITR and in different ATP states to provide molecular snapshots of ITR engagement, melting, and translocation. Additionally, our cryo-EM structure of Rep40 bound to the clinically relevant serotype AAV8 shows how Rep helicases dock onto the 5-fold of empty capsids in preparation for packaging. The motor-capsid structure also reveals a capsid ATP site that may influence capsid functions. Together, our structural data provide mechanistic insights into genome encapsidation of an important viral gene therapy vector.

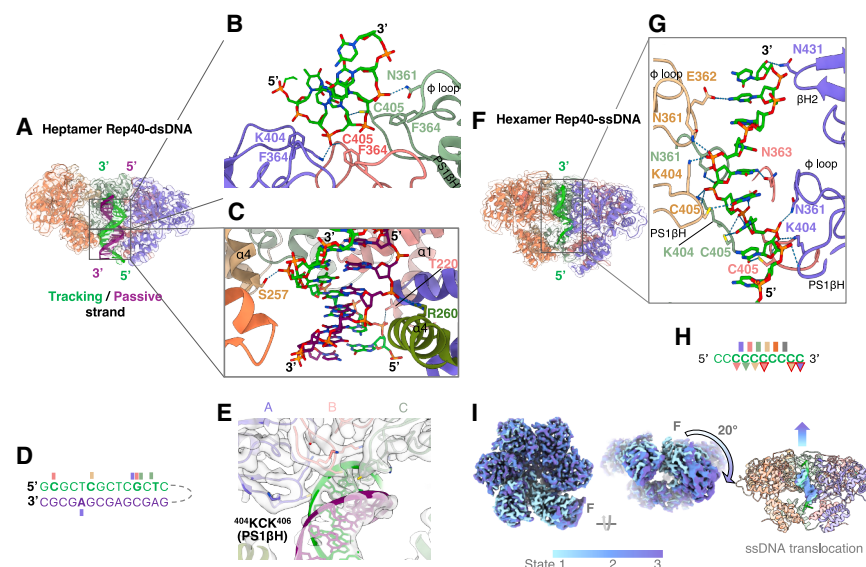
RESULTS

Rep40 helicase binds the hairpin ITR as heptamers and hexamers

To understand how Rep40 uses its oligomeric states to carry out helicase functions on the ITR packaging marker, we captured

and visualized complexes by cryo-EM. As AAV2 Rep40 forms transient oligomers,⁴² an N-terminal oligomerization linker from Rep78/68²⁷ (aa 201–224) was incorporated into our Rep40 construct and is hereby referred to as Rep40 ([Figures S1C](#) and [S1D](#)). Using mass photometry, we observed that Rep40 exists in several oligomeric states, ranging from dimer and trimer to hexamer, heptamer, and octamer in a concentration-dependent manner ([Figure S1E](#)). Notably, we did not detect tetrameric or pentameric species for Rep40. Cryo-EM analysis of Rep40 in complex with AAV2 5' ITR and ATP_γS showed Rep40 forms both heptamers and hexamers when bound to DNA ([Figures 1](#) and [S2](#)), similar to those observed for the Rep68 helicase domain.^{35,36} For all resolved Rep40-ITR complexes, the N-terminal oligomerization domain forms a star-shaped collar, while the AAA+ lobes assume a compact or relaxed conformation dependent on nucleotide-binding state ([Figure 1](#)). Additionally, consecutive nucleotide binding on adjacent Rep40 molecules induces compaction of subunits, resulting in asymmetric arrangements about the central DNA axis. Using 3D classification of images ([Figure S2](#)), we resolved structures of heptameric Rep40 bound to dsDNA at 2.8 Å ([Figures 1A–1C](#)) and hexameric Rep40 bound to ssDNA at 2.4 Å ([Figures 1D–1F](#); [Table S1](#)).

Our structures show how Rep's ATPase cycle allosterically affects gripping of the ITR-DNA. ATP binding triggers compaction of the AAA+ domain toward the lumen, which positions the pore loops (pre-sensor-1-β-hairpin [PS1βH] and φ-loop) for DNA binding. In the heptamer-dsDNA complex, the tracking strand is engaged by the pore loops of the three nucleotide-bound subunits (A, B, and C; [Figures 1B](#) and [1C](#)). The ATPase cycle starts when ATP binds to subunit A, where the Arg-finger (R444) of



(G) PS1βH and φ-loop of subunits A (purple), B (pink), C (light green), and D (tan) bind ssDNA (lime green). Interacting residues are shown as sticks.
(H) Schematic showing ssDNA that contacts specific subunits. Backbone interactions are indicated by rectangles; nucleotide base interactions are shown by triangles (subunits colored as in (G)).
(I) Superposition of three ssDNA-bound hexamer maps in different states of translocation.
See also Figures S2 and S4.

subunit B contacts the γ -phosphate to prepare it for hydrolysis (Figure S3A). This active state is maintained for one step in the cycle before hydrolysis occurs in the next subunit (subunit C). In contrast, in the hexamer, the ssDNA is engaged by the pore loops of all six nucleotide-bound subunits (Figures 1E and 1F). Here, the newly bound ATP is held for four steps before being hydrolyzed at the fifth step (subunit E; Figure S3B). Comparing the two assemblies, the ATPase cycle is shorter in the heptamer (three steps) than in the hexamer (six steps). These structures reveal how Rep40 tunes its oligomeric state to accommodate ds- and ssDNA regions of the ITR for melting and translocation.

Rep40 couples ATP hydrolysis to dsDNA distortion and ssDNA translocation via a pore-loop network

To understand how the ATPase cycle impacts Rep40 binding to ITR dsDNA, we investigated the DNA interactions in detail. In the heptamer, each of the three nucleotide-bound subunits provides a set of pore loops to engage dsDNA (Figures 2A and 2B). The three subunits' pore loops are held in a spiral staircase arrangement through inter-locking salt bridges positioning K404 and N361 for DNA backbone binding (Figures 2B, 2D, and S3C). Spacing between PS1βH and φ-loops is maintained by the bulky F364 (Figure 2B), a residue that has been shown to be essential for duplex melting by Rep68.³⁶ The pore-loop network also inserts C405 into the minor groove, which may contribute to duplex distortion during ATP hydrolysis (Figure 2E). Conversely, DNA contacts within the oligomerization domain are non-consecutive, involving 2 bp on the tracking strand and 1 bp on the passive strand (Figures 2C and 2D). In summary, our structure shows how heptameric Rep40 binds dsDNA and provides a model for how the pore-loop wedges deform the minor groove.

Figure 2. Rep helicase couples ATP hydrolysis with ITR distortion and translocation

(A) Cross-section of the map-model overlay for Rep40 heptamer in complex with dsDNA of the ITR and ATP γ S. DNA duplex with tracking (green) and passive (purple) strands colored. (B) DNA-binding loops PS1βH and φ-loop of subunits A (purple), B (pink), and C (light green) engage the tracking strand. Interacting residues are shown as sticks. (C) Oligomerization domain DNA interactions. Subunits colored (A, purple; B, pink; C, light green; D, tan; E, orange; G, dark green). Interacting residues are shown as sticks. (D) ITR schematic showing bases of the tracking (green) and passive (purple) strands that contact specific subunits (rectangles colored as in (B) and (C)). (E) Minor groove interaction with PS1βH loops of subunits A (purple), B (pink), and C (light green). (F) Cross-section of the map-model overlay for Rep40 hexamer in complex with ssDNA of the ITR (lime green) and ATP γ S.

By contrast, hexameric Rep40 forms a narrow lumen and uses all six sets of pore loops for ssDNA binding (Figures 2F and 2G). These pore loops form a right-handed helical arrangement along the tightly wound ssDNA backbone, forcing DNA base stacking and a 6-bp pitch. Cysteine residues within the pore loops (C405) directly bind DNA bases, consistent with a role in sensing DNA polarity. Subunit A also binds DNA via beta-hairpin-2 (βH2). The ssDNA interactions span 10 bp, forming extensive interactions with the backbone and base components of DNA (Figure 2H).

While these structures resolved discrete oligomeric states of the complex, helicases are dynamic molecular machines. To investigate the molecular mechanism of ssDNA translocation, we further classified images of the hexameric complex (Figure S4A). Comparing these states, we observe motions in which subunit F AAA+ lobe rotates $\sim 20^\circ$ upon releasing ADP (Figure 2I and S3D–E). This triggers the dissociation of DNA contacts at the 3' end, while ATP re-binding allows DNA binding at the subsequent 5' base pair (Figure S3F). This mechanism supports a translocation stoichiometry of one base pair per ATP hydrolyzed. Together, these structural analyses explain the mechanism by which hexameric Rep40 translocates ssDNA in a sequential and directional manner.

Putative mechanism for DNA melting by Rep40

We identified a sub-population ($\sim 7\%$) of hexameric Rep40 particles bound to double-stranded DNA in our cryo-EM dataset (Figures 3A and S4B). To resolve heterogeneous DNA features within this subgroup, we used 3D variability analyses to investigate molecular motions of the complex. Our analyses show that as the hexamer transitions from three to four compacted subunits, the duplex dsDNA undergoes melting resulting in

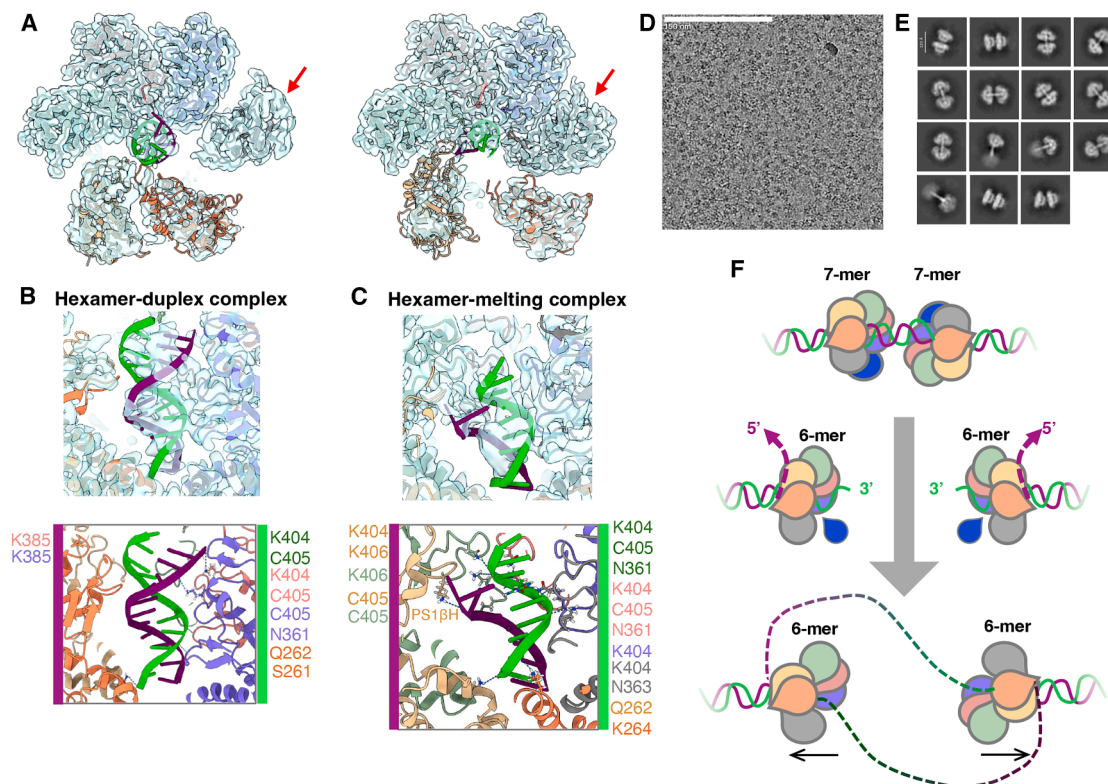


Figure 3. dsDNA melting mechanism by hexameric Rep

(A) Map-model overlay of the duplex dsDNA-bound (left) and forked DNA-bound (right) hexamers. View from top down, with the compaction of a 4th subunit (red arrow) correlating with duplex melting.

(B) Map-model overlay of the hexamer-duplex complex (top panel) and DNA-binding interactions (bottom panel). Interacting residues are shown as sticks. Residues (colored according to their subunit identity) and their interacting strands (tracking, lime green; passive, purple) are labeled along the periphery of the inset. Subunits A (purple), B (pink), C (light green), and E (orange) are shown.

(C) Map-model overlay of the hexamer-melting complex (top panel) and DNA-binding interactions (bottom panel). Interacting residues are shown as sticks. Residues (colored according to their subunit identity) and their interacting strands are shown. Subunits A (purple), B (pink), C (light green), D (tan), E (orange), F (gray), and G (dark green) are shown.

(D) Micrograph showing higher-order Rep40 oligomers on ITR. Scale bar is 150 nm in length.

(E) 2D class averages of double-stacked rings. Scale bar is 120 angstroms in length.

(F) Schematic of dsDNA melting mechanism. Heptamers load on dsDNA, the transition to hexameric Rep melts dsDNA, and hexameric Rep translocates ssDNA. See also [Figures S2](#) and [S4](#).

forked DNA ([Figures 3A–3C](#)). The duplex was best resolved in reconstructed frame 18 of 3D variability analyses and thus used for model fitting of the hexamer-duplex complex ([Figures 3B](#) and [S4C](#); [Table S1](#)). Compared to the ssDNA complex, the duplex-bound hexamer has a more constricted oligomerization domain. Consequently, B-form dsDNA is sterically strained.

3D variability analyses highlighted the complexity of Rep40 conformations bound to the ITR. We next used the output of these analyses to seed subsequent classification and refinement of this particle subgroup to solve the hexamer-melting complex ([Figures 3C](#) and [S4D](#); [Table S1](#)). Similar to the ssDNA complex, the tracking strand of DNA in the melting complex is tightly wound with stacked bases. To understand what triggers strand melting, we compared DNA interactions between the duplex- and melting-hexamer complexes. The number of subunits contacting the DNA increases from three to four upon unwinding. The four ATP-bound compacted subunits of the melting complex contribute an additional set of pore loops, which interact

with the tracking strand and may contribute to distortions that facilitate strand separation. Although the passive strand is not well resolved beyond the fork nexus, its proximity to subunit D's PS1 β H is consistent with guided exit from the helicase lumen ([Figure 3C](#)). To investigate if Rep40 forms higher-order oligomers on the ITR, Rep40-ITR particles were re-extracted from the raw micrographs with a larger box size ([Figure 3D](#)). 2D classification of these images revealed the presence of head-to-tail and head-to-head stacked Rep helicase rings ([Figure 3E](#)). Together, these snapshots of ITR-Rep40 assemblies reveal a multi-step melting mechanism in which (1) the heptamer docks on dsDNA and introduces minor distortions to the duplex, (2) loss of one subunit constricts the pore, (3) ATP-driven compaction of an additional nucleotide-bound subunit promotes duplex separation, and (4) the resulting hexamer drives fork migration through steric exclusion of the passive strand ([Figure 3F](#)). This progression underscores how increasing nucleotide occupancy dictates oligomer geometry to drive changes in oligomeric state and DNA binding.

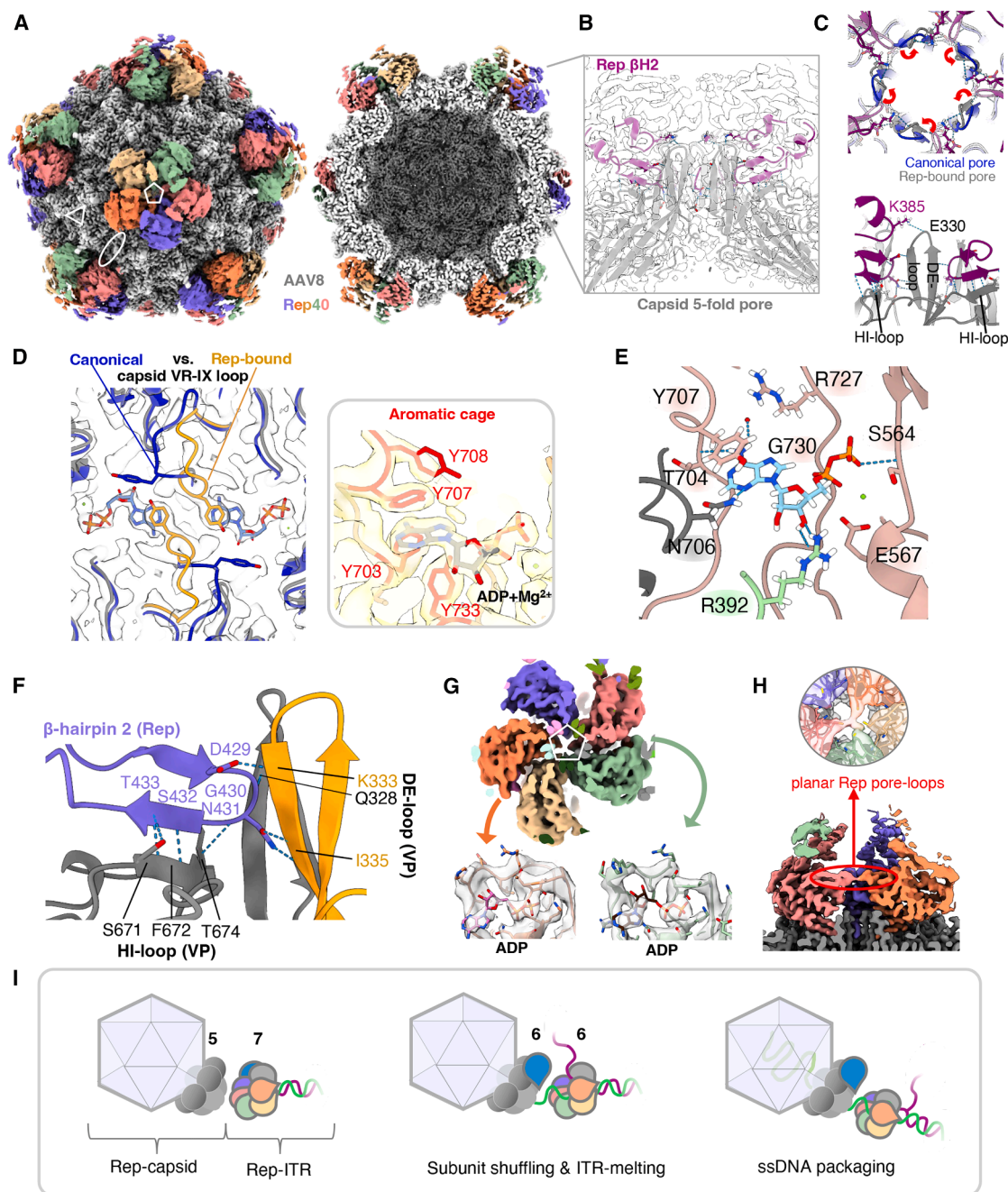


Figure 4. Pentameric Rep40 binds empty AAV8 capsid 5-fold pores

(A) Cryo-EM structure of the Rep40-AAV8 capsid complex. Icosahedral symmetry axes (5-, 3-, and 2-fold) are indicated by shapes. Cross section (right) shows Rep pentamers on capsid 5-fold pores. Rep40 oligomer colored by subunit (A, purple; B, pink; C, light green; D, tan; E, orange). AAV8 capsid is gray.

(B) Map-model overlay of the ordered Rep40 (purple, ribbons) and AAV8 VP3 (gray, ribbons) zoomed in on the 5-fold capsid pore.

(C) Superposition of AAV8 (dark blue; PDB ID: 6v12) with the Rep (purple)-bound AAV8 (gray) complex showing rotation of the AAV DE loop tips (indicated by red arrows). Lower panel shows residues involved in stabilizing the DE loop position in the Rep-bound structure.

(D) Same superposition as in (C) but highlighting the 2-fold axis ADP-binding site. The VR-IX loop residues are shown as sticks; the Rep-bound AAV8 VR-IX loop is highlighted in yellow (left panel). 2-fold VR-IX loop re-arrangement creates an aromatic cage (shown as sticks) in the Rep40-bound structure (right panel).

(E) Capsid VP3 residues that bind ADP and Mg^{2+} . The proposed Arg finger (R727) is highlighted.

(F) Residues involved in binding Rep40 β -hairpin-2 (purple). The hairpin binds the HI loop from one capsid monomer (VP, gray) and the DE loop from a second capsid monomer (yellow). Interacting residues are shown as sticks.

(G) Cryo-EM map of the locally refined, asymmetric Rep40-AAV8 pentameric complex. 5-fold symmetry axis is shown as a geometric shape. Insets show density for ADP between Rep40 dimers.

(legend continued on next page)

Rep40 helicase docks on AAV8 capsid 5-fold pore as a pentamer

To understand how Rep40 binds the capsid to deliver the ITR-marked genome, we used cryo-EM to solve the structure of Rep40 in complex with AAV8. We incubated Rep40 together with AAV8, ATP γ S, and mild crosslinker BS3 to stabilize the motor-capsid complex. Imposing icosahedral symmetry, we solved the structure of the complex to 1.9 Å (Figure 4; Table S2). We observe density corresponding to a Rep40 pentamer bound to each 5-fold vertex of the virus capsid (Figures 4A and S5). 3D classification of particles without symmetry showed that 94.4% of AAV8's 5-fold vertices are occupied by Rep40 (Figure S5). Recognizing that the AAV8 sample contains a mixture of empty and full capsids, we used 2D classification of images to separate capsids based on levels of interior DNA density. We show that while empty capsids have clear signal for Rep40, density for Rep40 was weak on DNA-loaded capsids (Figure S5). These data provide the first experimental evidence that Rep40 preferentially binds empty capsids.

Rep40 binds the empty AAV8 capsid via the core AAA+ domain. Lower resolution in the Rep40 oligomerization domain relative to the AAA+ domain within the symmetrized map indicates flexibility and asymmetry (Figures 4A and 4B). This suggests that, rather than the oligomerization domain, capsid binding drives Rep pentamerization. Specifically, five Rep β -hairpins (aa 423–445, β H2) insert between the capsid protein DE loops that form the 5-fold pore (Figure 4C). The inserted Rep40 β H2 is stabilized by H-bonds with the capsid DE loop and HI loop, resulting in an intermolecular anti-parallel β -sheet (Figures 4C and 4F). Rep binding also twists the capsid DE-loop tips (E330), resulting in constriction of the 5-fold pore entrance and widening of the base (Figures 4C and S6A). Consequently, interior electrostatics of the capsid are remodeled, exposing a basic patch at the rim, consistent with a role in DNA funneling (Figure S6B). Taken together, insertion of Rep's β H2 into the capsid 5-fold pore primes the capsid for packaging by widening the base of the pore and changing the interior electrostatics, while constricting the pore tips to deter nonspecific DNA threading.

Rep40-binding reveals a cryptic ATP-binding site on the empty capsid 2-fold

By comparing our structures of full AAV8 and Rep40-bound empty AAV8, we discovered a rearrangement of the variable region 9 (VR-IX; Figure S6C). The VR-IX loops on Rep40-bound empty AAV8 are “unbuckled,” flipping away from the 2-fold symmetry partner and weakening capsid boundaries (Figure 4D, left panel). In this arrangement, the phenol-ring of Y707 is stacked on top of an extra density, assigned as the adenine ring of an ADP. The adenine ring is nestled in an aromatic cage (Y733, Y703, Y707, and Y708), with Y707 acting as a lid by forming π -stacking interactions (Figure 4D, right panel). The adenine ring

is also H-bonded by T704 together with N706 from the opposing 2-fold partner (Figure 4E). The phosphates are coordinated by Mg $^{2+}$ and bound by G730 and S564. The Mg $^{2+}$ ion is also coordinated by E567 and S564, residues potentially involved in ATP hydrolysis (Figure 4E). Given its proximity to the ADP and dual side chain conformations, we suggest that R727 may act as an Arg finger, sensing γ -phosphates (Figures 4E and S6D). This putative ATP site exhibits features of Walker ATPases: aromatic cage, polar phosphate sensor (S564), acidic Mg $^{2+}$ coordinator (E567), and Arg finger (R727).^{43,44} In summary, our structure reveals a cryptic ATP-binding site on the 2-fold plateau of the Rep40-bound empty AAV8 capsid, suggesting that ATP may serve as an allosteric modulator of the capsid.

Capsid bound Rep40 pentamer adopts multiple conformations

To explore potential symmetry mismatches in Rep40 binding, we used symmetry expansion of our icosahedral reconstruction followed by asymmetric local refinement of 5-fold vertices (Figure S5). 3D classification of symmetry expanded particles showed that Rep40 pentamers were arranged asymmetrically on capsids outside of the inserted β H2 “pins” (Figure 4F). A continuous flexibility analysis revealed dynamic dimerization of pentamer subunits held together by ADP binding (Figure 4G). When viewed as a series, dimers switch sides like a paper fortune teller, with β H2 as pivot points anchored to the 5-fold pore (Video S1). In addition, there are minor movements in the protruded VR-I and VR-VI loops on the capsid influenced by its interaction with the Rep40 AAA+ domain. The Rep40 AAA+ lumen is sealed in these asymmetric pentamers. Further refinement of the best resolving class to 2.2 Å revealed planar arranged PS1 β H loops sealing across the capsid 5-fold pore (Figure 4H). The planar arrangement of the pore loops and absence of a continuous ATPase cycle suggest the pentameric form of Rep is not in DNA-binding mode. Together, Rep40 binding not only alters the shape of the 5-fold pore but also seals the entrance to prevent non-specific DNA packaging.

Rep40 and ADP binding triggers additional capsid interior rearrangements. Both nucleotide binding at the capsid 2-fold (VR-IX) and Rep binding at the 5-fold (HI loop) rearrange the β I strand underneath (Figures S6E and S6F). This shift affects the capsid C terminus and the VR-III- β F “wire” that directly binds the nucleotide (via R392; Figure 4E). The wire spans the 3-fold protrusions and base of the 5-fold pore. These interactions form an allosteric relay, coupling nucleotide occupancy and Rep binding to the structural state of the 5-fold packaging pore. These rearrangements affect neighboring pieces, communicating surface binding events across the capsid (Figure S6G). In summary, Rep40 binding alters the capsid pore size, seals the entrance and, together with nucleotide binding, triggers structural rearrangements to prime the pre-formed AAV capsids for packaging.

(H) Cryo-EM map of the locally refined, asymmetric Rep40-AAV8 vertex. Circled region and inset show sealing of the pore lumen by Rep40 PS1 β H loops. Rep40 colored by subunit and capsid colored as in (A).

(I) Schematic of DNA packaging by Rep40.

See also Figures S5–S7.

DISCUSSION

Our work captures multiple snapshots along the Rep-mediated DNA melting process and clarifies the later stages of the model proposed for Rep68.³⁶ Hexamers are the active form of many viral helicases for internal melting,^{45–47} steric extrusion driven fork migration and progressive translocation.^{48,49} “Docking” heptamer to “active” hexamer transitions have been reported for several other helicases including RuvB,⁵⁰ the bacteriophage T7 gp4,⁵¹ and mitochondrial Twinkle.⁵² By contrast, some viral helicases such as SV40 LTag,⁴⁵ papillomavirus E1,⁴⁹ and MpoxV E5⁴⁶ use double hexamers for origin melting. We also observe higher-order oligomers of Rep40 bound to ITR. Therefore, we propose head-to-head heptameric Rep helicases can locally melt dsDNA, allowing effective hexamer-driven strand separation in the absence of existing overhangs.

Our structures demonstrate the versatility of Rep40 in adopting distinct oligomeric states for its helicase and packaging roles. We resolved the first parvovirus motor-capsid complex, showing that Rep binds the empty capsid as a pentamer via β H2. The dependence of Rep pentamer formation on capsid 5-fold templating, rather than oligomerization “latch” linker,²⁷ may explain the high packaging efficiency of small Reps. While large Reps typically form hexameric or heptameric rings via “latch” linkers, small Reps are typically monomers in solution,^{23,53} free to assemble into the pentamers on empty capsids. In addition, the orientation of the pentameric Rep bound to AAV capsid is consistent with 3′ to 5′ directionality of packaging¹⁷ and agrees with mutagenesis studies showing the importance of the 5-fold pore in genome packaging.^{38,39} Based on our structure, a packaging efficient Rep mutant identified previously¹⁴ (Q439T) likely affects β H2-capsid interactions. We also show capsid VR-I protrusions potentially bind Rep, corroborating mutagenesis studies showing its importance for 5-fold pore dynamics and packaging.^{54,55} Furthermore, based on sequence and structural prediction analysis (Figure S7), we anticipate conservation of this Rep-capsid complex across the vertebrate parvoviruses, including human B19, bocavirus, and non-primate parvoviruses.

Taking our structures together, we propose an encapsidation model where the capsid-bound Rep pentamer serves as a docking platform for the ITR-bound heptamer (Figure 4I; Video S2). Upon association, the heptamer shuffles a subunit into the pentamer, transforming the inactive “cap” into a hexameric, capsid-anchored motor. Simultaneously, subunit loss in the ITR-bound heptamer triggers duplex melting and initiates ssDNA insertion. Consistent with our priming complex, we anticipate Rep oligomeric transitions to allosterically regulate capsid dynamics at packaging initiation. Viruses package a single copy of their genome by restricting DNA insertion to a single vertex. For icosahedral viruses without a unique portal protein such as AAV, all 5-fold pores are potential packaging sites. While Rep proteins dock indiscriminately across all twelve pores, we speculate that once the first active motor forms at a single pore, it allosterically inactivates the remaining pores. The symmetry mismatch between the hexameric Rep motor and the pentameric capsid pore parallels dsDNA viral packaging systems.^{56–58} Furthermore, Rep’s transition from a symmetric docking ring to an asymmetric spiral assembly on DNA mirrors the conformational

shifts of phage large terminase motors.⁵⁸ Critical questions remain regarding packaging initiation. Since small Rep proteins are promiscuous translocases and large Reps target both ITR arms, 3′ ITR selection for packaging initiation may involve host factors. Notable candidates for further investigation include FKBP52 and nucleophosmin.^{59,60} Regardless of the “initiation key,” a model of stochastic Rep priming and initiation, combined with allosteric control of the remaining pores, provides a plausible mechanism for single-site packaging regulation.

Finally, our discovery of a cryptic ATP-binding site on the Rep40-bound empty AAV8 capsid illustrates the genetic economy of this minimalistic virus. Previous studies established that capsid 2-fold residues are important for cell entry, pH sensing, endosomal escape, and transcriptional activation.^{61–64} In our structure, these same residues bind ADP and Mg^{2+} to regulate empty capsid dynamics during packaging. This highlights the 2-fold axis in coordinating various capsid activities required for viral fitness. Furthermore, the putative ATP-hydrolysis activity introduces an enzymatic perspective on these protein shells. Intriguingly, mutation of 2-fold residues showed no significant defects in packaging.⁶¹ As ATP concentration is also a regulator of phase separation,⁶⁵ it begs the question if AAV could exploit this for re-location between nucleolus/nucleoplasm and replication/transcription hubs. In conclusion, our structures provide insights into AAV biology and enable future structure-informed engineering to improve DNA packaging in gene therapy vectors.

Limitations of the study

In this work, the role of the Rep helicase domain in DNA melting, translocation, and empty capsid binding was examined. While the helicase domain is central to all four Rep proteins, this study does not account for the influence of the origin-binding or zinc-finger domains on these tasks. We comment on the functional advantage of small Reps in forming pentamers on capsids over large Reps; however, the distinct role of each Rep in packaging remains an open question. Furthermore, although this model integrates observations of DNA melting and asymmetric pentamer flexibility, the complete active packaging complex involving Rep, the ITR, and the capsid remains to be resolved.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for materials should be directed to and will be fulfilled by the lead contact, Sarah L. Rouse (s.rouse@imperial.ac.uk).

Materials availability

All unique/stable reagents generated in this study are available from the lead contact (s.rouse@imperial.ac.uk) with a completed materials transfer agreement.

Data and code availability

- Structural data generated in this study have been deposited in the electron microscopy databank (EMDB) and Protein Data Bank (PDB) under accession codes: EMD-54050 and PDB 9RM5 for hexameric Rep40 in complex with ssDNA ITR; EMD-54184 and PDB 9RQT for heptameric Rep40 in complex with dsDNA ITR; EMD-54416 and PDB 9SON for hexameric Rep40 dsDNA ITR melting structure; EMD-54429 (3DVA frame 18 map), EMD-54403 (consensus map), and PDB 9S10 for hexameric Rep40 dsDNA ITR duplex structure; EMD-54195 and PDB 9RRS for

whole capsid Rep40-emptyAAV8 complex; and EMDB-54328 and PDB 9RWG for local asymmetric structure of Rep40-AAV8 5-fold pore.

- Raw cryo-electron microscopy dataset of the AAV2Rep40+AAV8 capsid+ATPγS has been deposited at EMPIAR as EMPIAR-12986-.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

ACKNOWLEDGMENTS

The research in the S.L.R. group is supported by UKRI Future Leaders Fellowship (MR/Y01975X/1). We thank Suhail Islam for computational support and Paul Simpson for EM support at Imperial College London Centre for Structural Biology. Cryo-EM data processing used a BBSRC-funded GPU server (BB/X019284/1). We thank Diamond for access and support of the cryo-EM facilities at the UK national electron bio-imaging centre (eBIC): the Rep40+ITR dataset was collected from session B133230-24 on Krios I, and the Rep40+AAV8 capsid dataset from B139228-2 on Krios III. We acknowledge the following for helpful discussion: Chris Binny, Altar Munis, Natalie Tigue, Lisa Vinal, Kyran Sandu, Davide Busato, Julie Douthwaite, and members of the Rouse and Buebeck groups. We thank Meng (Cleo) Zhang for creation of the animated movie (Video S2). We thank Rhian Jones and Christian Speck for critical reading of the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, V.X., M.W., J.D.B., D.B., and S.L.R.; methodology, V.X., A.M., J.D.B., D.B., and S.L.R.; investigation, V.X., J.D.B., and M.A.-G.-d.-E.; writing – original draft, V.X., D.B., and S.L.R.; writing – review & editing, V.X., J.D.B., D.B., and S.L.R.; funding acquisition, D.B. and S.L.R.; resources, J.D.B., M.W., M.A.-G.-d.-E., D.B., and S.L.R.

DECLARATION OF INTERESTS

M.W. and M.A.-G.-d.-E. are employees of AstraZeneca. V.X. is supported by a PhD studentship part-funded by AstraZeneca to S.L.R.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

No generative AI was used in this work.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
 - Bacterial strains
 - Cell lines
- **METHOD DETAILS**
 - Construct design
 - Protein expression and purification
 - AAV production
 - Cryo-EM sample preparation
 - Cryo-EM data acquisition
 - Cryo-EM processing and model fitting
 - Sequence alignments and structural model predictions
 - Mass photometry
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2026.117044>.

Received: November 10, 2025

Revised: January 5, 2026

Accepted: February 4, 2026

REFERENCES

1. Li, C., and Samulski, R.J. (2020). Engineering adeno-associated virus vectors for gene therapy. *Nat. Rev. Genet.* 21, 255–272. <https://doi.org/10.1038/s41576-019-0205-4>.
2. Gao, G., Vandenberghe, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X., and Wilson, J.M. (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>.
3. Calcedo, R., Vandenberghe, L.H., Gao, G., Lin, J., and Wilson, J.M. (2009). Worldwide epidemiology of neutralizing antibodies to adeno-associated viruses. *J. Infect. Dis.* 199, 381–390. <https://doi.org/10.1086/595830>.
4. Wang, J.H., Gessler, D.J., Zhan, W., Gallagher, T.L., and Gao, G. (2024). Adeno-associated virus as a delivery vector for gene therapy of human diseases. *Signal Transduct. Target. Ther.* 9, 78. <https://doi.org/10.1038/s41392-024-01780-w>.
5. Mével, M., Bouzelha, M., Leray, A., Pacouret, S., Guilbaud, M., Penaud-Budloo, M., Alvarez-Dorta, D., Dubreil, L., Gouin, S.G., Combail, J.P., et al. (2020). Chemical modification of the adeno-associated virus capsid to improve gene delivery. *Chem. Sci.* 11, 1122–1131. <https://doi.org/10.1039/C9SC04189C>.
6. Bryant, D.H., Bashir, A., Sinai, S., Jain, N.K., Ogden, P.J., Riley, P.F., Church, G.M., Colwell, L.J., and Kelsic, E.D. (2021). Deep diversification of an AAV capsid protein by machine learning. *Nat. Biotechnol.* 39, 691–696. <https://doi.org/10.1038/s41587-020-00793-4>.
7. Martin, J., Frederick, A., Luo, Y., Jackson, R., Joubert, M., Sol, B., Poulin, F., Pastor, E., Armentano, D., Wadsworth, S., and Vincent, K. (2013). Generation and Characterization of Adeno-Associated Virus Producer Cell Lines for Research and Preclinical Vector Production. *Hum. Gene Ther. Methods* 24, 253–269. <https://doi.org/10.1089/HGTB.2013.046>.
8. Jalsić, L., Lytvyn, V., Elahi, S.M., Hrapovic, S., Nassoury, N., Chahal, P.S., Gaillet, B., and Gilbert, R. (2023). Inducible HEK293 AAV packaging cell lines expressing Rep proteins. *Mol. Ther. Methods Clin. Dev.* 30, 259–275. <https://doi.org/10.1016/j.OMTM.2023.07.002>.
9. Spisak, S., Ndhairo, N., Betenbaugh, M.J., and Ostermeier, M. (2025). Deep Mutational Scanning of the AAV rep Gene to Assess Effects on DNA Packaging in Expi293F Suspension Culture. *Biotechnol. Bioeng.* 122, 3061–3072. <https://doi.org/10.1002/BIT.70041>.
10. Shitik, E.M., Shalik, I.K., and Yudkin, D.V. (2023). AAV- based vector improvements unrelated to capsid protein modification. *Front. Med.* 10, 1106085. <https://doi.org/10.3389/FMED.2023.1106085>.
11. Satkunathan, S., Thorpe, R., and Zhao, Y. (2017). The function of DNA binding protein nucleophosmin in AAV replication. *Virology* 510, 46–54. <https://doi.org/10.1016/J.VIROL.2017.07.007>.
12. Faust, S.M., Bell, P., Cutler, B.J., Ashley, S.N., Zhu, Y., Rabinowitz, J.E., and Wilson, J.M. (2013). CpG-depleted adeno-associated virus vectors evade immune detection. *J. Clin. Investig.* 123, 2994–3001. <https://doi.org/10.1172/JCI68205>.
13. Kosaka, M., Sajiki, A.F., Fujita, K., Yamada, K., Kawano, K., Hirazawa, K., Matsuno, T., Takada, M., Shimozawa, N., Inoue, K.I., et al. (2025). Evaluation of the loading capacity and patterns of packaged DNA in AAV genomes of different sizes using long-read sequencing. *Mol. Ther. Methods Clin. Dev.* 33, 101474. <https://doi.org/10.1016/J.OMTM.2025.101474>.
14. Jain, N.K., Ogden, P.J., and Church, G.M. (2024). Comprehensive mutagenesis maps the effect of all single codon mutations in the AAV2 rep gene on AAV production. *eLife* 12, RP87730. <https://doi.org/10.7554/eLife.87730.2>.
15. Steininger, T., Öttl, V., Franken, L.E., Frank, C., Ohland, P., Lopez Ferreira, M., Klostermann, S., Fritsch, J., Hirschauer, E., Sandmeier, A., et al. (2025).

- Improved Recombinant Adeno-Associated Viral Vector Production via Molecular Evolution of the Viral Rep Protein. *Int. J. Mol. Sci.* 26, 1319. <https://doi.org/10.3390/ijms26031319>.
16. Azim, T., Myerscough, D., and Silberg, J.J. (2024). Synonymous mutations in AAV Rep enhance genome packaging in a library selection. *eLife* 14, RP105040. <https://doi.org/10.1101/2024.10.08.617207>.
17. King, J.A., Dubielzig, R., Grimm, D., and Kleinschmidt, J.A. (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J.* 20, 3282–3291. <https://doi.org/10.1093/EMBOJ/20.12.3282>.
18. Wang, X.S., Ponnazhagan, S., and Srivastava, A. (1996). Rescue and replication of adeno-associated virus type 2 as well as vector DNA sequences from recombinant plasmids containing deletions in the viral inverted terminal repeats: selective encapsidation of viral genomes in progeny virions. *J. Virol.* 70, 1668–1677. <https://doi.org/10.1128/jvi.70.3.1668-1677.1996>.
19. Snyder, R.O., Im, D.S., Ni, T., Xiao, X., Samulski, R.J., and Muzyczka, N. (1993). Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. *J. Virol.* 67, 6096–6104. <https://doi.org/10.1128/JVI.67.10.6096-6104.1993>.
20. McCarty, D.M., Pereira, D.J., Zolotukhin, I., Zhou, X., Ryan, J.H., and Muzyczka, N. (1994). Identification of linear DNA sequences that specifically bind the adeno-associated virus Rep protein. *J. Virol.* 68, 4988–4997. <https://doi.org/10.1128/JVI.68.8.4988-4997.1994>.
21. Ling, C., Wang, Y., Lu, Y., Wang, L., Jayandharan, G.R., Aslanidi, G.V., Li, B., Cheng, B., Ma, W., Lentz, T., et al. (2015). The Adeno-Associated Virus Genome Packaging Puzzle. *J. Mol. Genet. Med.* 9, 175. <https://doi.org/10.4172/1747-0862.1000175>.
22. Cao, M., You, H., and Hermonat, P.L. (2014). The X gene of Adeno-Associated Virus 2 (AAV2) is involved in viral DNA replication. *PLoS One* 9, 104596. <https://doi.org/10.1371/journal.pone.0104596>.
23. Smith, R.H., and Kotin, R.M. (1998). The Rep52 Gene Product of Adeno-Associated Virus Is a DNA Helicase with 3'-to-5' Polarity. *J. Virol.* 72, 4874–4881. <https://doi.org/10.1128/JVI.72.6.4874-4881.1998>.
24. Han, S.-I., Kawano, M.-A., Ishizu, K.-I., Watanabe, H., Hasegawa, M., Kanesashi, S.-N., Kim, Y.-S., Nakanishi, A., Kataoka, K., and Handa, H. (2004). Rep68 protein of adeno-associated virus type 2 interacts with 14-3-3 proteins depending on phosphorylation at serine 535. *Virology* 320, 144–155. <https://doi.org/10.1016/j.virol.2003.11.024>.
25. Cassinotti, P., Weitzand, M., and Tratschin, J.D. (1988). Organization of the adeno-associated virus (AAV) capsid gene: Mapping of a minor spliced mRNA coding for virus capsid protein. *Virology* 167, 176–184. [https://doi.org/10.1016/0042-6822\(88\)90067-0](https://doi.org/10.1016/0042-6822(88)90067-0).
26. Dubielzig, R., King, J.A., Weger, S., Kern, A., and Kleinschmidt, J.A. (1999). Adeno-Associated Virus Type 2 Protein Interactions: Formation of Pre-Encapsidation Complexes. *J. Virol.* 73, 8989–8998. <https://doi.org/10.1128/JVI.73.11.8989-8998.1999>.
27. Zarate-Perez, F., Bardelli, M., Burgner, J.W., Villamil-Jarauta, M., Das, K., Kekilli, D., Mansilla-Soto, J., Linden, R.M., and Escalante, C.R. (2012). The Interdomain Linker of AAV-2 Rep68 Is an Integral Part of Its Oligomerization Domain: Role of a Conserved SF3 Helicase Residue in Oligomerization. *PLoS Pathog.* 8, e1002764. <https://doi.org/10.1371/journal.ppat.1002764>.
28. Di Pasquale, G., and Stacey, S.N. (1998). Adeno-Associated Virus Rep78 Protein Interacts with Protein Kinase A and Its Homolog PRKX and Inhibits CREB-Dependent Transcriptional Activation. *J. Virol.* 72, 7916–7925. <https://doi.org/10.1128/jvi.72.10.7916-7925.1998>.
29. Grifman, M., Chen, N.N., Gao, G.P., Cathomen, T., Wilson, J.M., and Weitzman, M.D. (1999). Overexpression of Cyclin A Inhibits Augmentation of Recombinant Adeno-Associated Virus Transduction by the Adenovirus E4orf6 Protein. *J. Virol.* 73, 10010–10019. <https://doi.org/10.1128/JVI.73.12.10010-10019.1999>.
30. Nash, K., Chen, W., Salganik, M., and Muzyczka, N. (2009). Identification of Cellular Proteins That Interact with the Adeno-Associated Virus Rep Protein. *J. Virol.* 83, 454–469. <https://doi.org/10.1128/JVI.01939-08>.
31. Nash, K., Chen, W., and Muzyczka, N. (2008). Complete in vitro reconstitution of adeno-associated virus DNA replication requires the minichromosome maintenance complex proteins. *J. Virol.* 82, 1458–1464. <https://doi.org/10.1128/JVI.01968-07>.
32. Nash, K., Chen, W., McDonald, W.F., Zhou, X., and Muzyczka, N. (2007). Purification of host cell enzymes involved in adeno-associated virus DNA replication. *J. Virol.* 81, 5777–5787. <https://doi.org/10.1128/JVI.02651-06>.
33. Srivastava, A., Lusby, E.W., and Berns, K.I. (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>.
34. Prasad, K.M., Zhou, C., and Trempe, J.P. (1997). Characterization of the Rep78/Adeno-Associated Virus Complex. *Virology* 229, 183–192. <https://doi.org/10.1006/VIRO.1996.8431>.
35. Santosh, V., Musayev, F.N., Jaiswal, R., Zárate-Pérez, F., Vandewinkel, B., Dierckx, C., Endicott, M., Sharifi, K., Dryden, K., Henckaerts, E., and Escalante, C.R. (2020). The Cryo-EM structure of AAV2 Rep68 in complex with ssDNA reveals a malleable AAA+ machine that can switch between oligomeric states. *Nucleic Acids Res.* 48, 12983–12999. <https://doi.org/10.1093/nar/gkaa1133>.
36. Jaiswal, R., Braud, B., Hernandez-Ramirez, K.C., Santosh, V., Washington, A., and Escalante, C.R. (2025). Cryo-EM structure of AAV2 Rep68 bound to integration site AAVS1: insights into the mechanism of DNA melting. *Nucleic Acids Res.* 53, gkaf033. <https://doi.org/10.1093/nar/gkaf033>.
37. Hickman, A.B., Ronning, D.R., Perez, Z.N., Kotin, R.M., and Dyda, F. (2004). The Nuclease Domain of Adeno-Associated Virus Rep Coordinates Replication Initiation Using Two Distinct DNA Recognition Interfaces. *Mol. Cell* 13, 403–414. [https://doi.org/10.1016/S1097-2765\(04\)00023-1](https://doi.org/10.1016/S1097-2765(04)00023-1).
38. Bleker, S., Sonntag, F., and Kleinschmidt, J.A. (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>.
39. Bleker, S., Pawlita, M., and Kleinschmidt, J.A. (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>.
40. James, J.A., Escalante, C.R., Yoon-Robarts, M., Edwards, T.A., Linden, R.M., and Aggarwal, A.K. (2003). Crystal Structure of the SF3 Helicase from Adeno-Associated Virus Type 2. *Structure* 11, 1025–1035. [https://doi.org/10.1016/S0969-2126\(03\)00152-7](https://doi.org/10.1016/S0969-2126(03)00152-7).
41. James, J.A., Aggarwal, A.K., Linden, R.M., and Escalante, C.R. (2004). Structure of adeno-associated virus type 2 Rep40–ADP complex: Insight into nucleotide recognition and catalysis by superfamily 3 helicases. *Proc. Natl. Acad. Sci. USA* 101, 12455–12460. <https://doi.org/10.1073/pnas.0403454101>.
42. Dignam, S.S., Collaco, R.F., Bieszczad, J., Needham, P., Trempe, J.P., and Dignam, J.D. (2007). Coupled ATP and DNA binding of adeno-associated virus Rep40 helicase. *Biochemistry* 46, 568–576. <https://doi.org/10.1021/BI061762V>.
43. Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1, 945–951. <https://doi.org/10.1002/J.1460-2075.1982.TB01276.X>.
44. Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop — a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* 15, 430–434. [https://doi.org/10.1016/0968-0004\(90\)90281-F](https://doi.org/10.1016/0968-0004(90)90281-F).
45. Shahid, T., Danazumi, A.U., Tehseen, M., Alhudhali, L., Clark, A.R., Savva, C.G., Hamdan, S.M., and De Biasio, A. (2025). Structural dynamics of DNA

- p>unwinding by a replicative helicase.
- Nature*
- 641, 240–249.
- <https://doi.org/10.1038/s41586-025-08766-w>
- .
46. Zhang, W., Liu, Y., Yang, M., Yang, J., Shao, Z., Gao, Y., Jiang, X., Cui, R., Zhang, Y., Zhao, X., et al. (2024). Structural and functional insights into the helicase protein E5 of Mpox virus. *Cell Discov.* 10, 67. <https://doi.org/10.1038/s41421-024-00680-1>.
 47. Cheng, Y., Han, P., Peng, Q., Liu, R., Liu, H., Yuan, B., Zhao, Y., Kuai, L., Qi, J., Miao, K., et al. (2025). Assembly and breakage of head-to-head double hexamer reveals mpox virus E5-catalyzed DNA unwinding initiation. *Nat. Commun.* 16, 5176. <https://doi.org/10.1038/s41467-025-60539-1>.
 48. Zhang, Y., Fan, B., Gao, Y., Yang, J., Zhang, W., Su, S., Li, L., Li, H., Luo, Z., Tang, G., et al. (2025). Structural and functional studies of the main replication protein NS1 of human parvovirus B19. *Nucleic Acids Res.* 53, 562. <https://doi.org/10.1093/NAR/GKAF562>.
 49. Schuck, S., and Stenlund, A. (2015). A Conserved Regulatory Module at the C Terminus of the Papillomavirus E1 Helicase Domain Controls E1 Helicase Assembly. *J. Virol.* 89, 1129–1142. <https://doi.org/10.1128/jvi.01903-14>.
 50. Miyata, T., Yamada, K., Iwasaki, H., Shinagawa, H., Morikawa, K., and Mayanagi, K. (2000). Two different oligomeric states of the RuvB branch migration motor protein as revealed by electron microscopy. *J. Struct. Biol.* 131, 83–89. <https://doi.org/10.1006/JSBI.2000.4290>.
 51. Satapathy, A.K., Kochaniak, A.B., Mukherjee, S., Crampton, D.J., Van Oijen, A., and Richardson, C.C. (2010). Residues in the central β -hairpin of the DNA helicase of bacteriophage T7 are important in DNA unwinding. *Proc. Natl. Acad. Sci. USA* 107, 6782–6787. <https://doi.org/10.1073/pnas.1002734107>.
 52. Fernández-Millán, P., Lázaro, M., Cansız-Arda, Ş., Gerhold, J.M., Rajala, N., Schmitz, C.A., Silva-España, C., Gil, D., Bernadó, P., Valle, M., et al. (2015). The hexameric structure of the human mitochondrial replicative helicase Twinkle. *Nucleic Acids Res.* 43, 4284–4295. <https://doi.org/10.1093/nar/gkv189>.
 53. Yoon-Roberts, M., Blouin, A.G., Bleker, S., Kleinschmidt, J.A., Aggarwal, A.K., Escalante, C.R., and Linden, R.M. (2004). Residues within the B' motif are critical for DNA binding by the superfamily 3 helicase Rep40 of adeno-associated virus type 2. *J. Biol. Chem.* 279, 50472–50481. <https://doi.org/10.1074/jbc.M403900200>.
 54. Asokan, A., Hamra, J.B., Govindasamy, L., Agbandje-McKenna, M., and Samulski, R.J. (2006). Adeno-Associated Virus Type 2 Contains an Integrin $\alpha 5 \beta 1$ Binding Domain Essential for Viral Cell Entry. *J. Virol.* 80, 8943–8950. <https://doi.org/10.1128/jvi.00843-06>.
 55. Drouin, L.M., Lins, B., Janssen, M., Bennett, A., Chipman, P., McKenna, R., Chen, W., Muzyczka, N., Cardone, G., Baker, T.S., and 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>.
 56. Liu, Y.T., Jih, J., Dai, X., Bi, G.Q., and Zhou, Z.H. (2019). Cryo-EM structures of herpes simplex virus type 1 portal vertex and packaged genome. *Nature* 570, 257–261. <https://doi.org/10.1038/s41586-019-1248-6>.
 57. Fang, Q., Tang, W.C., Tao, P., Mahalingam, M., Fokine, A., Rossmann, M.G., and Rao, V.B. (2020). Structural morphing in a symmetry-mismatched viral vertex. *Nat. Commun.* 11, 1713. <https://doi.org/10.1038/s41467-020-15575-4>.
 58. Pajak, J., Dill, E., Reyes-Aldrete, E., White, M.A., Kelch, B.A., Jardine, P.J., Arya, G., and Morais, M.C. (2021). Atomistic basis of force generation, translocation, and coordination in a viral genome packaging motor. *Nucleic Acids Res.* 49, 6474–6488. <https://doi.org/10.1093/NAR/GKAB372>.
 59. Bevington, J.M., Needham, P.G., Verrill, K.C., Collaco, R.F., Basrur, V., and Trempe, J.P. (2007). Adeno-associated virus interactions with B23/ Nucleophosmin: identification of sub-nucleolar virion regions. *Virology* 357, 102–113. <https://doi.org/10.1016/J.VIROL.2006.07.050>.
 60. Zhao, W., Zhong, L., Wu, J., Chen, L., Qing, K., Weigel-Kelley, K.A., Larsen, S.H., Shou, W., Warrington, K.H., and Srivastava, A. (2006). Role of cellular FKBP52 protein in intracellular trafficking of recombinant adeno-associated virus 2 vectors. *Virology* 353, 283–293. <https://doi.org/10.1016/J.VIROL.2006.04.042>.
 61. Aydemir, F., Salganik, M., Resztak, J., Singh, J., Bennett, A., Agbandje-McKenna, M., and Muzyczka, N. (2016). Mutants at the 2-Fold Interface of Adeno-associated Virus Type 2 (AAV2) Structural Proteins Suggest a Role in Viral Transcription for AAV Capsids. *J. Virol.* 90, 7196–7204. <https://doi.org/10.1128/jvi.00493-16>.
 62. Lochrie, M.A., Tatsuno, G.P., Christie, B., McDonnell, J.W., Zhou, S., Surosky, R., Pierce, G.F., and Colosi, P. (2006). Mutations on the external surfaces of adeno-associated virus type 2 capsids that affect transduction and neutralization. *J. Virol.* 80, 821–834. <https://doi.org/10.1128/JVI.80.2.821-834.2006>.
 63. Salganik, M., Aydemir, F., Nam, H.-J., McKenna, R., Agbandje-McKenna, M., and Muzyczka, N. (2014). Adeno-Associated Virus Capsid Proteins May Play a Role in Transcription and Second-Strand Synthesis of Recombinant Genomes. *J. Virol.* 88, 1071–1079. <https://doi.org/10.1128/JVI.02093-13>.
 64. Nam, H.-J., Gurda, B.L., McKenna, R., Potter, M., Byrne, B., Salganik, M., Muzyczka, N., and Agbandje-McKenna, M. (2011). Structural studies of adeno-associated virus serotype 8 capsid transitions associated with endosomal trafficking. *J. Virol.* 85, 11791–11799. <https://doi.org/10.1128/JVI.05305-11>.
 65. Patel, A., Malinowska, L., Saha, S., Wang, J., Alberti, S., Krishnan, Y., and Hyman, A.A. (2017). ATP as a biological hydrotrope. *Science* 356, 753–756. <https://doi.org/10.1126/science.aaf6846>.
 66. Bepler, T., Kelley, K., Noble, A.J., and Berger, B. (2020). Topaz-Denoise: general deep denoising models for cryoEM and cryoET. *Nat. Commun.* 11, 5208. <https://doi.org/10.1038/s41467-020-18952-1>.
 67. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* 75, 861–877. <https://doi.org/10.1107/S2059798319011471>.
 68. Croll, T.I. (2018). ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* 74, 519–530. <https://doi.org/10.1107/S2059798318002425>.
 69. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501. <https://doi.org/10.1107/S0907444910007493>.
 70. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. <https://doi.org/10.1002/pro.3943>.
 71. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>.
 72. van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M., Söding, J., and Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* 42, 243–246. <https://doi.org/10.1038/s41587-023-01773-0>.
 73. Dutta, T., Sasidharan, K., Ciociola, E., Pennisi, G., Noto, F.R., Kovoovu, L., Kroon, T., Lindblom, A., Du, Y., Pirmoradian, M., et al. (2024). Mitochondrial amidoxime-reducing component 1 p.Ala165Thr increases protein degradation mediated by the proteasome. *Liver Int.* 44, 1219–1232. <https://doi.org/10.1111/liv.15857>.
 74. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). Cryo-SPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* 14, 290–296. <https://doi.org/10.1038/nmeth.4169>.

75. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J., and Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* 16, 1153–1160. <https://doi.org/10.1038/s41592-019-0575-8>.
76. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 493–500. <https://doi.org/10.1038/s41586-024-07487-w>.
77. Mietzsch, M., Barnes, C., Hull, J.A., Chipman, P., Xie, J., Bhattacharya, N., Sousa, D., McKenna, R., Gao, G., and 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>.
78. Durrant, J.D. (2019). Blendmol: Advanced macromolecular visualization in Blender. *Bioinformatics* 35, 2323–2325. <https://doi.org/10.1093/bioinformatics/bty968>.
79. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. <https://doi.org/10.1038/msb.2011.75>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
LEMO21(DE3) E.coli cells	New England Biolabs	Cat#C2528J
AAV8 capsids	This paper	N/A
Chemicals, peptides, and recombinant proteins		
IPTG	Sigma-Aldrich	Cat#I6758
Benzonase endonuclease	Merck Millipore	Cat#70746
Imidazole	Sigma-Aldrich	Cat#I5513
BS3 cross-linker	Thermo Fisher	Cat#A39266
ATP _γ S	Abcam	Cat#ab146714
MgCl ₂	Sigma-Aldrich	Cat#M8266
TCEP	Thermo Fisher	Cat#77720
Glycerol	Sigma-Aldrich	Cat#G5516
Freestyle F17 media	Gibco	Cat#A1383501
Polyethyleneimine	PolySciences	Cat#24765
GlutaMAX	Gibco	Cat#35050061
Polaxamer 188	Gibco	Cat#24040032
NeutrAvidin protein	Thermo Fisher	Cat#31000
Mycoalert mycoplasma detection	Lonza	Cat# LT07-318
Deposited data		
cryoTEM dataset of AAV2 Rep40 + AAV8 capsid + ATP _γ S	This paper	EMPIAR-12986
cryoEM model and map of hexameric Rep40 in complex with ITR-ssDNA and ATP _γ S-ADP-Mg ²⁺	This paper	PDB entry ID: 9RM5 ; EMD-54050
cryoEM model and map of heptameric Rep40 in complex with ITR-dsDNA and ATP _γ S-ADP-Mg ²⁺	This paper	PDB entry ID: 9RQT ; EMD-54184 (composite map), EMD-54181 (consensus map), EMD-54182 (3-mer focused map), EMD-54183 (4-mer focused map)
cryoEM model and map of hexameric Rep40 in complex with duplex ITR-dsDNA and ATP _γ S-Mg ²⁺ (Duplex complex)	This paper	PDB entry ID: 9S10 ; EMD-54429 (3DVA series 0 frame 18 map), EMD-54403 (consensus map)
cryoEM model and map of hexameric Rep40 in complex with forked ITR-DNA and ATP _γ S-Mg ²⁺ (Melting complex)	This paper	PDB entry ID: 9SON ; EMD-54416
cryoEM model and map of Rep40-bound AAV8 empty capsid in complex with ADP-Mg ²⁺ (Whole capsid map)	This paper	PDB entry ID: 9RRS ; EMD-54195
cryoEM model and map of Rep40-bound AAV8 5-fold vertex in complex with ADP-Mg ²⁺ (Local map)	This paper	PDB entry ID: 9RWG ; EMD-54328 (local map), EMD-54195 (consensus map)
Experimental models: Cell lines		
Human: VPC2.0 HEK293F	Thermo Fisher	#CatA49784
Oligonucleotides		
AAV2 5' ITR (flip configuration)	GenScript Biotech	GenBank: OP161118.1
Recombinant DNA		
pET21a(+)-Rep40-FLAG-His	This paper	N/A
pAAV Rep-Cap	This paper	N/A
Helper plasmid (AAVpro system)	TaKaRa	#Cat623066526655
Software and algorithms		
EPU (v3.8.1)	FEI, Thermo Fisher Scientific	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
cryoSPARC (v4.4.1)	Structura Biotechnology	https://cryosparc.com
Topaz (v0.2.5)	Bepler et al. ⁶⁶	https://github.com/tbepler/topaz
Phenix (v1.21.2)	Liebschner et al. ⁶⁷	https://phenix-online.org
ISOLDE (v1.9)	Croll et al. ⁶⁸	https://tristanic.github.io/isolde/
COOT (v0.9.8.96)	Emsley et al. ⁶⁹	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
UCSF ChimeraX (v1.8)	Pettersen et al. ⁷⁰	https://www.cgl.ucsf.edu/chimerax/
Blender	Blender Foundation	https://www.blender.org
Jalview (v2)	Waterhouse et al. ⁷¹	https://www.jalview.org
Clustal Omega	EMBL-EBI	https://www.ebi.ac.uk/jdispatcher/msa/clustalo
FoldSeek	van Kempen et al. ⁷²	https://search.foldseek.com
DiscoverMP	Refeyn	https://refeyn.com/software-release-updates
Other		
UltrAuFoil R1.2/1.3 grids	Quantifoil	Cat#UF1213300
HisTrap HP column	Cytiva	Cat#17524801
HiTrap Q HP column	Cytiva	Cat#17115401
HiTrap SP HP column	Cytiva	Cat#29051324
Superose 6 Increase 10/300 GL	Cytiva	Cat#29091596
Slide-A-Lyzer Dialysis Cassette	Thermo Fisher	Cat#66380
PES protein concentrators (100 kDa MWCO)	Millipore	Cat#UFC910024

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial strains

Escherichia coli strain LEMO21(DE3) (NEB) was cultured in LB medium with 37°C for plasmid transformation and protein expression.

Cell lines

VPC2.0 cells (Thermo Fisher, A49784) were cultured in Freestyle F17 medium (Gibco) supplemented with 20 mM GlutaMAX and 0.001% Poloxamer 188 at 37°C in suspension for AAV packaging. Cells were authenticated using STR profiling and monitored for mycoplasma contamination using Lonza Mycoalert mycoplasma detection kit (Cat# LT07-318).

METHOD DETAILS

Construct design

Sequence encoding AAV2 Rep40 with an extra N-terminal sequence from Rep78/68 included for oligomer stabilisation (Rep68 residues 201–224²⁷; Uniprot- P03132), equivalent to Rep68 Δ1–200, followed by a thrombin site, FLAGx3 tag and TEV cleavage site were cloned into a pET21a(+) backbone with 6xHis tag expression and purchased from GenScript Biotech. The AAV2 5' ITR (GenBank accession: OP161118.1) was synthesised by GenScript Biotech.

Protein expression and purification

pET21a(+)-Rep40-FLAG-His plasmids were transformed into LEMO21(DE3) cells (NEB) and single colonies were grown at 37°C until OD600 reaches 0.5 where 0.2 mM IPTG was added and the culture subsequently incubated at 18°C for overnight protein expression. Cells were harvested the following day by centrifugation at 4500 g at 4°C for 20 min and resuspended with 50 mL/L pellet of pre-cooled cell lysis buffer (20 mM Tris, 500 mM NaCl, 5 mM imidazole, 1 mM TCEP, 10% Glycerol and protease inhibitor; pH 7.4 at 4°C), then sonicated at 48% amplitude with 9 s-on 9 s-off cycle while kept on ice. Benzonase endonuclease was added to lysate per manufacturer's instructions and incubated (10 min at 37°C) before centrifuging at 4°C, 35000 g for 30 min. The pellet was discarded and supernatant was used for Ni-IMAC, AIEX and SEC purification (Figure S1D). Cell lysates were applied to a HisTrap HP column equilibrated in buffer (20 mM Tris, 500 mM NaCl, 5 mM imidazole, 1 mM TCEP, 10% glycerol; pH 7.4, 4°C). Bound protein was washed with low-imidazole buffer (20 mM Tris, 500 mM NaCl, 1 mM TCEP, 10% glycerol and 20 mM imidazole; pH 7.4 at 4°C) and eluted with high-imidazole buffer (all components identical except 300 mM imidazole). The eluate was diluted in imidazole-free buffer and subjected to AIEX on a HiTrap Q HP column using a NaCl gradient (100 mM–1 M; 20 mM Tris, 1 mM TCEP, 1 mM MgCl₂). Protein-containing fractions were concentrated with a 30 kDa MWCO spin concentrator (Amicon) and further purified by SEC on a

Superose 6 Increase 10/300 GL column equilibrated in buffer (25 mM Tris, 200 mM NaCl, 1 mM TCEP, 1 mM MgCl₂, 1% glycerol; pH 8.0, room temperature). Peak fractions were pooled, concentrated with 30 kDa centrifugal concentrators, and protein concentration was determined by Nanodrop (Thermo Fisher).

AAV production

The AAV genome contained human MARC1 (hMARC1) A165 driven by the ApoE/haAT1 promoter.⁷³ This chimeric promoter contains the core promoter from the human alpha-1-antitrypsin (haAT) gene and the enhancer/hepatic locus control region from the apolipoprotein E (ApoE) gene. Packaging of this genome into AAV8 capsids was performed using the co-transfection of three plasmids into VPC2.0 cells using Polyethyleneimine reagent (DNA: PEI = 1:2). The plasmids for transfection were at 1:1:1 mass ratio with the genome, packaging (pAAV Rep-Cap) and helper plasmid (AAVpro system). VPC2.0 cells were grown in Freestyle F17 supplemented with 20mM GlutaMAX and 0.001% Polaxamer 188. Packaged AAV was purified by ion-exchange chromatography using HiTrap SP HP column. Briefly, transfected cells at 72 h post transfection were pelleted and sonicated with the pulse 1 second on and off at 35% amplitude, followed by Benzonase-based treatment (Merck). The treated cells were centrifuged to remove precipitated contaminants, and the supernatant was purified for AAV according to the manufacturer's instructions. The eluted AAV was concentrated using PES protein concentrators (100 kDa MWCO) using formulation buffer containing glycerol and Plutonic F-68 in Dulbecco's Phosphate Buffered Saline.

Cryo-EM sample preparation

For Rep40-ITR-ATPγS samples: 18 μM of Rep40-FLAG-His protein in buffer (25 mM Tris, 200 mM NaCl, 1 mM TCEP, 1% glycerol, 1mM MgCl₂ pH 7.8) was incubated with 1 mM ATPγS and 2 μM AAV2 5' ITR DNA at room temperature. For Rep40-AAV8Capsid-ATPγS samples: 36 μM Rep40-FLAG-His, and AAV8 (1E¹³ vg/mL) was incubated 1:1 v/v then 1 mM ATPγS was added and the mixture incubated for 5 min before dialyzing in 1L buffer (25mM HEPES, 200mM NaCl, 1% glycerol, 1 mM MgCl₂ pH 7.8) using a mini Slide-A-lyzer (20 kDa MWCO) on 20°C hotplate for 30 min. BS3 cross-linker was then added per manufacturer's instructions, for a further 2 h incubation before making grids.

UltrAuFoil R1.2/1.3 grids were glow discharged for 90 s in air at 70% power using a NanoClean 1070 (Fischione Instruments). Vitrobot Mark IV plunge-freezer (Thermo Fisher) was used for grid freezing with chamber kept at 100% humidity and 25°C, the wait-time was set at 10s with 8s blot time, 0 blot force. 3 μL of sample were deposited on to grids, plunge-frozen in liquid ethane with setting described and stored in LN₂.

Cryo-EM data acquisition

For Rep40-ITR-ATPγS-Mg²⁺ dataset: 8,500 Movies were collected on a 300 kV Titan Krios microscope (X-FEG, Thermo Fisher) equipped with Gatan K3 direct electron detector, BioContinuum energy filter (20 eV slit), 70 μm C2 aperture and 100 μm objective aperture at ×105,000 magnification (0.825 Å/pix) with EPU software 3.8.1 (AFIS/FFI) at electron Biology Imaging Center eBIC (Didcot, Oxfordshire, UK). Individual exposures were acquired at a dose rate of 16.2 e⁻/pix/s at a total dose of 39.9 e⁻/Å. The nominal defocus was between range −0.6 μm and −2.4 μm in 0.3 μm increments.

For Rep40-AAV8 dataset: 17,295 Movies were collected on a 300 kV Titan Krios microscope (X-FEG, Thermo Fisher) equipped with Falcon 4i direct electron detector and TFS selectris X energy filter at ×130,000 magnification (0.931 Å/pix) with EPU software 3.8.1 (AFIS/FFI) at electron Biology Imaging Center eBIC (Didcot, Oxfordshire, UK). Individual exposures were acquired at a dose rate of 7.2 e⁻/pix/s at a total dose of 40 e⁻/Å. The nominal defocus was between range −0.6 μm and −2.4 μm in 0.3 μm increments.

Cryo-EM processing and model fitting

Movie stacks were imported into cryoSPARC⁷⁴ v4. 4. 1 for processing, beam-induced motion was corrected by patch motion correction and CTF was estimated by patch CTF estimation.

For Rep40-ITR-ATPγS-Mg²⁺ dataset (Figures S2 and S4): 8,500 Movies were imported and a subset of micrographs (381) were used for initial particle picking using blob picker, the extracted particles were used to generate 2D templates for Topaz (v 0.2.5)⁷⁵ training. The Topaz model was used to extract 4,416,824 particles from the entire dataset. Particles were subjected to 2D classification and selection to yield 2,265,855 particles which was used for *Ab initio* reconstruction (5 class). The classes were refined with hetero refinement then non-uniform (NU) refinement to 2.77 Å for the hexamer-ssDNA structure and 3.05 Å for the heptamer dsDNA structure. The hexamer structure is then subjected to *Ab initio* reconstruction (3 class) where 1 class (319,129 particles) was taken forward for further 3D classification (10 class) and 1 resulting class (class 3) taken forward for NU-refinement followed by reference-based motion correction, CTF refinement and local refinement with a soft mask around the whole protein to 2.44 Å (FSC 0.143, from now on unless stated otherwise) "HexRep40-ssDNA-ATPγS-ADP-Mg²⁺". The 3.05 Å heptamer dsDNA structure from early NU-refinement was further 3D classified to 10 classes and regrouped to 4 classes where 1 class (138,238 particles) was further NU-refined to 3.07 Å. It was then subjected to reference-based motion correction, CTF and local refinement with a soft mask covering the whole protein to 2.75 Å. Masks covering either half of the heptamer (3/7mer or 4/7mer) was generated from the consensus map and separately used for signal subtraction and masked local refinement to yield a 4-mer map (2.95 Å) and 3-mer map (2.63 Å), the 2 maps were stitched together to generate a composite map in Phenix 1.21.2 "HeptRep40-dsDNA-ATPγS-ADP-Mg²⁺". An earlier Hexamer class from the heterogeneous refinement (584,527 particles) was subjected to NU-refinement to 2.86 Å. The particles were then

3D classified and used for separate workflows: a) “Moving lobe hexamer-ssDNA workflow” (Figures S4A and S4B) “Hexamer-dsDNA workflow” (Figure S4B). A) All 10 classes were subjected to heterogeneous reconstruction and NU-refinement, 1 class (59,253 particles) was selected for further 3D classification (10 class), re-group (5-class) then NU-refined, resulting in 2 classes showing translocation states “state 2” (3.31 Å) and “state 3” (3.39 Å). B) 3D classes were re-grouped into 4 classes and NU-refined, resulting in 1 class (108,620 particles, 3.07 Å) bound to dsDNA. 3D variation (with 6 Å low-pass-filter) was used to produce 3 series of movies (20 frames each) where series 0 frame 1, 16, and 18 represented different stages of duplex melting. Frame 18 map was used to fit HexRep40-dsDNA-ATPγS-Mg²⁺ “Duplex complex” (Figure S4C). Frame 1, 16 and previous consensus volume (108,620 particles, 3.07 Å) was used as input volumes to 3D classify the larger set of particles from upstream processing (584,527 particles), after NU-refinement the HexRep40-dsDNA-ATPγS-Mg²⁺ “Melting complex” was solved at 2.92 Å (Figure S4D). To obtain 2D classes that showcase higher order oligomeric states, 13,635 particles were extracted using larger (360 pix) box size and 2D classified.

For Rep40-AAV8-ATPγS-Mg²⁺ dataset (Figure S5): 17,295 Movies were imported and templates from a previous screening dataset were used for template picking of AAV capsid particles generating 128,171 particles. 2D classification and clean up resulted in 128,100 particles. Ab initio reconstruction (I-symmetry) of 3 classes resulted in 125,855 empty Rep40-bound capsids and 2,200 full AAV8 capsids. NU-refinement of full capsids with I-symmetry imposed resulted in 2.27 Å reconstruction. NU-refinement (I-symmetry, symmetry relaxation by maximization) of the Rep40-bound empty capsids resulted in 2.33 Å map which was improved to 1.88 Å by CTF refinements and reference-based motion correction, resulting in the “Rep40-AAV8cap-ADP-Mg²⁺” map. Local refinement of the 5-fold vertex was performed by I-symmetry expansion of the capsid particles, particle subtraction with mask preserving only 1/12 5-fold vertices and 3D classification in C1 into 10 classes with 6 Å low-pass filter. Of the resulting classes, 5.6% (83,989 particles) were not bound by Rep40 and the rest has different asymmetrically arranged Rep40-pentamers bound. One class with 191,887 particles and 2.19 Å resolution was selected for model fitting to produce “PentRep40-5foldAAV8-ADP-Mg²⁺” model. Model fitting was done in ISOLDE⁶⁸ using Alpha Fold 3⁷⁶ predicted Rep40 models and AAV8 (PDB = 6v12⁷⁷) as starting models. Waters for the whole capsid model were fitted in COOT (v0.9.8.96).⁶⁹ Models were refined using ISOLDE v1.9 followed by Phenix⁶⁷ v1.21.2 real space auto-refinement and iterated. Visualisations used ChimeraX⁷⁰ and Blender.⁷⁸

Sequence alignments and structural model predictions

Multiple sequence alignment of AAV Rep40 serotypes used Clustal Omega,⁷⁹ followed by visualisation in Jalview (v2).⁷¹ Structural homologues of AAV2 Rep40 were identified with FoldSeek.⁷² Structural predictions of helicase-capsid interactions for these homologues used AlphaFold 3.⁷⁶

Mass photometry

The mass of purified Rep40 protein in solution was measured using TwoMP Auto (Refeyn). The mass photometer laser was warmed up for an hour before experiment. The objective was cleaned and sample cassette slide prepared in a dust-free manner. The cassette was then aligned and measurements calibrated using Neutravidin. 36 μM Rep40-FLAG-His stock in buffer (25 mM Tris, 200 mM NaCl, 1 mM TCEP, 1% glycerol, 1mM MgCl₂ pH 7.8) was pre-diluted to appropriate concentrations in PBS in a 96-well V-bottom plate. Experiment was run in auto-mode. The Data was plotted using the DiscoverMP software.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of the percentage of hexamers bound to dsDNA compared to ssDNA was done by dividing the number of particles belonging to the 3D classified group of dsDNA-bound hexamers (108,620) by the total number of hexamers (1,356,067) to yield 7.4%.

The ~20 (22) degrees angular rotation of Rep40's AAA+ domain upon ADP-release indicated in Figure 2I was obtained by overlaying the subunits and defining the angle between identical atoms on the center of the rotated AAA+ domains (aa 418 Ca) and a stationary point on the oligomerisation domain (aa 260 Ca).

Quantification of the percentage of 5-fold vertexes on the capsid that are occupied by Rep pentamers was done by dividing the number of occupied vertices (1,420,436) by the total number of vertices (1,504,452) to yield 94.4%. The number of vertices were obtained by treating each 5-fold vertex from the capsid as individual particles and 3D classifying to sort particles based on Rep-occupancy.