Localization of the N-terminus of minor coat protein IIIa in the adenovirus capsid

Running title: Localization of adenovirus IIIa N-terminus

Carmen San Martín [1*], Joel N. Glasgow [2,3,4], Anton Borovjagin [2], Matthew S. Beatty [2], Elena A. Kashentseva [2], David T. Curiel [2,4*], Roberto Marabini [5], Igor P. Dmitriev [2,4]

[1] Department of Macromolecular Structure, Centro Nacional de Biotecnología (CNB-CSIC), Darwin 3, 28049 Madrid, Spain

[2] Division of Human Gene Therapy, Departments of Medicine, Obstetrics and Gynecology, Pathology and Surgery, [3] Division of Cardiovascular Disease, [4] the Gene Therapy Center, University of Alabama at Birmingham, Birmingham, AL 35294, USA.

[5] Escuela Politécnica Superior, Universidad Autónoma de Madrid, Francisco Tomás y Valiente 11, 28049 Madrid, Spain

* Corresponding authors:

Carmen San Martín.

Centro Nacional de Biotecnología (CNB-CSIC), Darwin 3, 28049 Madrid (Spain)

Ph: 34 91 585 5450, Fax 34 91 585 4506, e-mail: carmen@cnb.csic.es

David T. Curiel, M.D., Ph.D.

Ph: (205) 934-8627, Fax: (205) 975-7476, e-mail: curiel@uab.edu

Summary

Minor coat protein IIIa is conserved in all adenoviruses and required for correct viral assembly, but its precise function in capsid organization is unknown. The latest adenovirus capsid model proposes that IIIa is located underneath the vertex region. To obtain experimental evidence on the location of IIIa and further define its role, we engineered the IIIa gene to encode heterologous N-terminal peptide extensions. Recombinant adenovirus variants with IIIa encoding six-histidine tag (6-His), 6-His and FLAG peptides, or 6-His linked to FLAG with a $(Gly_4Ser)_3$ linker were rescued and analyzed for virus yield, capsid incorporation of heterologous peptides, and capsid stability. Longer extensions could not be rescued. Western blot analysis confirmed that the modified IIIa proteins were expressed in infected cells and incorporated into virions. In the adenovirus encoding the 6-His-linker-FLAG-IIIa gene, the 6-His tag was present in light particles but not in mature virions. Immuno-electron microscopy of this virus showed that the FLAG epitope is not accessible to antibodies on the viral particles. Three-dimensional electron microscopy (3DEM) and difference mapping located the IIIa N-terminal extension beneath the vertex complex, wedged at the interface between penton base and the peripentonal hexons, therefore supporting the latest proposed model. The position of the IIIa N-terminus and its low tolerance for modification provide new clues for understanding the role of this minor coat protein in adenovirus capsid assembly and disassembly.

Keywords: Adenovirus, polypeptide IIIa, minor coat protein, virus structure, three-dimensional electron microscopy

Introduction

Adenovirus (Ad) has been found in every class of vertebrates including mammals, birds, fish, amphibians and reptiles [1]. In humans they cause common respiratory, gastrointestinal and eye infections, with 51 different serotypes described so far. Although the most common type of adenovirus infection is subclinical, adenovirus-induced diseases are responsible for significant morbidity in immuno-compromised patients and children in underdeveloped areas [2]. They are widely used as vectors for gene therapy, and have showed promise in vaccine delivery and oncolysis [3; 4]. Because of this dual relationship between adenoviruses and humans (as pathogenic agents and therapeutic tools), accurate knowledge of capsid structure is fundamental to both the discovery of anti-adenovirus drugs and the design of new, efficient viral vectors for gene therapy.

At least 11 structural proteins combine to form the non-enveloped icosahedral adenovirus capsid that contains the dsDNA-protein core complex. The virion has a diameter of approximately 900 Å not counting the fibers, and a particle mass of 150×10^6 Da [5]. Each capsid facet has 12 trimers of the major coat protein, hexon. At each vertex, five penton base subunits form the penton base, which is complexed with a trimer of the long projecting fiber. The crystal structures of hexon and the vertex proteins (penton base and fiber) have been solved [6; 7; 8; 9].

Atomic resolution structures are not presently available for any of the adenovirus minor coat proteins (polypeptides IIIa, VI, VIII and IX). These proteins are thought to play a dual role in stabilizing the virion while allowing the flexibility necessary for its disassembly during infection [10]. However, the specific role of each minor coat protein in capsid assembly is not known.

Polypeptide IIIa is conserved in all adenoviruses described so far [1], with an estimated copy number of 60 [5; 11; 12]. In human adenovirus type 5 (Ad5) the 585 amino acid precursor pIIIa is

cleaved by the viral protease during particle maturation, removing 15 residues from its C-terminus and giving rise to the 570 amino acid, 63.5 kDa mature protein [12; 13]. Polypeptide IIIa has a role in capsid assembly, as shown by studies on the Ad2 *ts112* mutant. When grown at the non-permissive temperature, *ts112* contains similar amounts of IIIa as the wild type virus, but it produces mainly low density particles depleted in core components [14]. These are reminiscent of assembly intermediates blocked at a late stage of assembly, probably DNA encapsidation.

Early experimental work produced somewhat conflicting evidence about the location of IIIa. On one hand, co-immunoprecipitation and cross-linking assays indicated that IIIa was interacting with core polypeptide VII, and was thus probably internally placed [12; 15]. On the other hand, a serum raised against IIIa was able to immunoprecipitate intact virions [15], suggesting exposure on the outer capsid surface. Assays estimating the degradation rates of proteins in the virion by Argon plasma etching found that IIIa and polypeptide VI are degraded at intermediate rates, suggesting a position inside the capsid but not deep in the core [16]. A model based on combined electron microscopy (EM)/X-ray imaging studies on Ad2 [11] fulfilled both requirements by placing IIIa at the facet edges, with about 2/3 of the protein bulk on the external capsid surface, while the rest was assigned to an inner density region directly below. However, recent studies have challenged this and other assignments of minor coat protein localizations in the capsid.

First, reconstructions of polypeptide IX-deleted mutants showed that the lack of IX resulted in loss of density at both the positions assigned to IX and the external part of IIIa [17; 18]. While Fabry and coworkers interpreted their maps in terms of IIIa being present only outside the capsid and destabilized in the absence of IX, Scheres and coworkers showed immunolabeling data suggesting that IIIa is not exposed at the virion surface. Second, a 3DEM map of adenovirus with GFP fused to the C-terminus of IX showed extra density above the external position assigned to

IIIa, raising the possibility that either the density on the capsid surface assigned to IIIa corresponds to the C-terminal region of IX instead, or that modifications to IX result in the ejection of IIIa [19]. Third, a 6 Å resolution adenovirus map interpreted with the help of secondary structure predictions has proposed new locations for all minor coat proteins, reassigning IIIa to an internal position close to the vertex [20]. Whereas all these interpretations are plausible and consistent with their respective data, the need for direct experimental evidence for minor capsid protein locations is clear. A recent study on the structure of canine adenovirus supports the location of the C-terminal domain of IX at the capsid edges [21], but there is at present no experimental confirmation regarding the position of IIIa or any of the other minor coat proteins.

Here we describe the design, construction and characterization of Ad5 mutants containing heterologous peptide extensions at the N-terminus of protein IIIa. The behavior of the different constructs regarding viral rescue, extension incorporation and capsid stability is evaluated. 3DEM analyses reveal the location of the IIIa N-terminal extension within the capsid. The implications of our findings for the role of polypeptide IIIa in viral assembly and disassembly are discussed.

Results

To investigate the role of polypeptide IIIa in adenovirus capsid assembly, we constructed recombinant Ad genomes encoding heterologous peptide extensions at the N-terminus of IIIa **(Fig. 1)**. The N-terminus was selected because the pIIIa C–terminus is cleaved off by the adenoviral protease during virion maturation [13]. Peptide extensions were incorporated into IIIa between the first two methionines and included a 6-His tag (Ad5H_6IIIa), an 18-residue extension comprised of a 6-His tag and a flexible FLAG-containing linker (Ad5Luc1-HFpIIIa), and a 31-residue extension containing a 6-His tag, a flexible linker containing three G_4S repeats and a FLAG peptide (Ad5GLflagIIIa). To explore the capacity of the IIIa capsid locale to tolerate more structurally complex moieties, we introduced a 60-residue extension containing a 37-residue (45Å) alpha helix spacer [22] between the 6-His and FLAG sequences (Ad5Luc1-H45FpIIIa). We have previously incorporated fluorescent proteins to other capsid components thereby allowing direct imaging of vector particles [23; 24]. To explore the compatibility and function of this moiety at the IIIa locale, we inserted the eGFP open reading frame between the FLAG and 6-His tags (Ad5Luc1-HeGFPFpIIIa).

Following construction and verification of recombinant Ad genomes containing modified IIIa, *Pac*I-linearized genomes were transfected into 293 cells for virus rescue. Three of the genomes, Ad5H_6IIIa, Ad5Luc1-HFpIIIa, and Ad5GLflagIIIa produced viral propagation and cytopathic effect (CPE). However, introduction of Ad5Luc1-H45FpIIIa and Ad5Luc1-HeGFPFpIIIa genomes into 293 cells did not result in CPE after multiple attempts. On this basis, we concluded that these IIIa species were not compatible with the formation of new viral particles and represented a block to Ad propagation.

The rescued Ad5H_6IIIa, Ad5Luc1-HFpIIIa, and Ad5GLflagIIIa constructs were propagated in 293 cells and analyzed for capsid incorporation of heterologous peptide sequences. The expression of modified IIIa genes in the context of Ad5H_6IIIa, Ad5Luc1-HFpIIIa, and Ad5GLflagIIIa genomes was validated by Western blot using monoclonal antibodies (mAb) specific for 6-His or FLAG peptides. As shown in **Fig. 2A**, multiple polypeptides containing 6-His or FLAG tag were observed in the lysates of 293 cells infected with Ad5H_6IIIa, Ad5Luc1-HFpIIIa, or Ad5GLflagIIIa virus but not with control Ad5GL vector. The detection of major protein bands with molecular mass of 65 kDa, as expected for the modified polypeptides IIIa, confirmed the synthesis of full-size protein IIIa containing the N-terminal FLAG and/or 6-His tags upon Ad infection. To study incorporation of modified IIIa proteins into the capsid, Ad5H_6IIIa, Ad5Luc1-HFpIIIa, and Ad5GLflagIIIa viruses were purified on CsCl gradients. Both the upper viral band containing defective viral particles and the lower band representing mature virions separated during the first centrifugation were collected and purified on a second CsCl gradient. We have not observed significant differences between virus preparations in amount of defective particles relative to mature virions. The total yields of mature virions purified from equal number of cells infected with Ad5H_6IIIa, Ad5Luc1-HFpIIIa, Ad5GLflagIIIa, and Ad5GL were estimated as 3.4×10^{12}, 1.7×10^{12}, 3.9×10^{12}, and 5.8×10^{12} viral particles, respectively. The capsid incorporation of modified IIIa polypeptides was assessed in the context of defective and normal virions using Western blot analysis for the presence of IIIa-incorporated 6-His and FLAG peptides as described above. As can be seen in **Fig. 2B,** the 6-His tag was present in Ad5H_6IIIa defective particles, while both the 6-His and FLAG tags were detected in immature Ad5Luc1-HFpIIIa and Ad5GLflagIIIa virions as was expected from the IIIa gene modifications introduced into the corresponding viral genomes. When mature virions were analyzed, positive

signals were detected for both Ad5Luc1-HFpIIIa and Ad5GLflagIIIa when probed with an anti-FLAG mAb, as well as for Ad5H_6IIIa and Ad5Luc1-HFpIIIa with an anti-Penta-His mAb. However, no 6-His-containing protein was detected in mature Ad5GLflagIIIa virions. While confirming the presence of intact 6-aa and 18-aa peptide extensions at the IIIa N-terminus in mature Ad5H_6IIIa and Ad5Luc1-HFpIIIa viral particles, these data indicated that a part of the 31-aa extension that includes the 6-His tag is absent from Ad5GLflagIIIa particles, even if the complete extension is correctly expressed and incorporated into immature particles.

Since previous evidence indicated a role for polypeptide IIIa in capsid assembly, we tested the thermostability of Ad5Luc1-HFpIIIa and Ad5GLflagIIIa virions to see whether extensions of the IIIa N terminus affected the efficiency of virus infection. The relative stability of Ad5Luc1-HFpIIIa, Ad5GLflagIIIa, and Ad5GL was assessed based on reduction of Ad-mediated gene transfer levels subsequent to virus incubation at 45°C which were calculated with respect to untreated virus samples. These experiments did not reveal any significant changes in infectivity of Ad5Luc1-HFpIIIa and Ad5GLflagIIIa viruses as compared to the Ad5GL control **(Fig. 3)** thereby indicating that the N-terminal extensions introduced into protein IIIa do not interfere with viral capsid integrity required for cell infection.

We then used Ad5GLflagIIIa for studies aimed to elucidate the location of the N-terminal IIIa modification in the capsid. We first performed immuno-electron microscopy (immuno-EM) using a gold-conjugated secondary Ab for visual detection of the N-terminal tag. As positive controls, we included Ad5EZ-pIX-Flag-45A-6His F5/3 and Ad5EZ-pIX-Flag-10A-6His F3/3 vectors that encode modified IX proteins displaying both FLAG and 6-His sequences separated by either 45Å or 10Å alpha helix spacers, respectively. No significant label was detected on Ad5GLflagIIIa viral particles, neither intact nor broken, in experiments performed with either

anti-6-His–tag or anti-FLAG antibodies. Both antibodies successfully recognized the corresponding epitopes when displayed at the polypeptide IX C-terminus (**Fig. 4**), even showing quantitative label differences related to the different accessibility of the 6-His tag when two constructions with different linker lengths were compared (**Table 1**). This indicates that the FLAG epitope in the IIIa extension, while present in the capsid as shown by Western blot, is masked by other viral components.

To localize the N-terminus of polypeptide IIIa in the capsid, we calculated a difference map between three-dimensional cryoEM reconstructions of the modified Ad5GLflagIIIa and of the original Ad5GL vector encoding native IIIa. Special care was taken to use similar imaging conditions for both datasets, thereby minimizing differences caused by the effect of the microscope contrast transfer function. Both maps were calculated using a similar number of images and the same image processing procedures, and filtered at the same final resolution (14 Å). Since the only difference between both specimens is the N-terminal extension in IIIa, positive difference peaks in an Ad5GLflagIIIa-Ad5GL difference map should reveal the position of the extra residues, and therefore of polypeptide IIIa N-terminus. **Fig. 5a** shows a central cross section of the Ad5GL and Ad5GLflagIIIa maps, and of the difference between them. The strongest positive difference peaks (over 3σ above noise level) appear inside the protein shell, close to the vertices. Vertex density was weaker than that of hexons in the Ad5GLflagIIIa map, which originated a negative difference peak encompassing the penton. The Ad5GLflagIIIa penton occupancy was only 45% that of the control Ad5GL. This indicates that on average, more than 6 pentons are lost per particle. It was therefore necessary to investigate the possibility that the positive difference peaks close to the penton arose as a consequence of structural rearrangements due to the absence of some vertex components.

Since it is not possible to differentiate particles lacking part of their pentons directly in the experimental images, we resorted to a single particle reconstruction (SPR)-based classification method [18]. SPRs were obtained for all particles in the Ad5GLflagIIIa dataset, using an algorithm previously shown to reduce the SPR artifacts at the symmetry axes. Two subsets containing approximately 30% of the complete original image dataset each were selected, and used to calculate two 3D maps corresponding to those particles with highest (Ad5GLflagIIIa_high) and lowest (Ad5GLflagIIIa_low) penton densities. These maps were subsequently compared to a map calculated from a similar number of Ad5GL particles selected at random throughout the defocus groups (Ad5GL_subset). Penton occupancy was 71% and 7% for the Ad5GLflagIIIa_high and Ad5GLflagIIIa_low maps, respectively. Difference maps calculated between either Ad5GLflagIIIa_high or Ad5GLflagIIIa_low, and the Ad5GL_subset map, at 18 Å resolution, are shown in **Fig. 5c**. The (Ad5GLflagIIIa_high minus Ad5GL_subset) map was noisier than the others, but it showed the same positive difference located adjacent to the innermost region of penton. A difference map between Ad5GLflagIIIa_high and Ad5GLflagIIIa_low showed no significant difference at this position (not shown). Therefore we conclude that the positive difference peak is not related to penton loss but to the presence of the IIIa N-terminal extension. The (Ad5GLflagIIIa_low minus Ad5GL_subset) map did not show any significant negative difference peaks other than the penton itself, indicating that even in the lowest penton occupancy particles no other icosahedrally ordered components were lost**.**

A surface rendering of the 14 Å resolution difference map contoured at 3σ sigma level over the Ad5GL 3D map is shown in **Fig. 6a, c**. There were no significant difference peaks outside the capsid. The strongest peaks above noise level are located on the inner shell surface, close to the interface between penton base and the peripentonal hexons. There are 60 such peaks in the

difference map, in agreement with the biochemically determined polypeptide IIIa copy number [11; 12]. The difference peak is nestled within a canyon (**Fig. 6c**), and is seen to be masked by other densities when the contouring threshold of the Ad5GL map is slightly lowered (not shown)**.** This may explain why there was no label in broken particles in our immuno-EM studies. Fitting of a FLAG peptide crystallographic model into one of the difference peaks shows that its volume is just large enough to contain the FLAG tag (**Fig. 6b**). Fitting of the high resolution structures of penton base [7] and hexon [6] onto the Ad5GL map provides a clearer view of the position of the IIIa N-terminus with respect to these capsid components (**Fig. 6d, e**). The IIIa N-terminal extension is located at the opening of a crevice defined by the penton base and two adjacent peripentonal hexons. The center of the difference peak is approximately 15 Å away from the closest atoms present in the penton base and hexon structures. Five residues in each terminus of hexon, and 51 residues at the penton base N-terminus, are not present in their crystal structures, but are in regions oriented towards the core. The N-terminus of polypeptide IIIa, as located by our difference mapping, is within reach of interactions with some of these terminal residues.

Discussion

We have tested the effect of modifying adenovirus minor coat protein IIIa on capsid assembly and stability. The N-terminus of IIIa was chosen for incorporating extensions, due to proteolysis of the C-terminus during virus maturation. We have determined the position of the N-terminus of adenovirus minor coat protein IIIa by peptide-based difference mapping of cryoEM 3D reconstructions. The power of this technique was first demonstrated when the N-terminus of hepatitis B virus capsid protein was localized by addition of eight extraneous residues, which were visible in the difference map even at moderate (20 Å) resolution [25]. Our results emphasize once more the potential of this methodology, by locating a similarly small extension in the much larger adenovirus capsid (~950 vs. ~300Å diameter, 20 times larger in mass).

The Ad5GLflagIIIa recombinant virus encodes a 31-residue extension at the N-terminus of IIIa. Western blot analyses showed that the entire 31 residue-modified IIIa protein was expressed in infected 293 cells and incorporated into low density Ad5GLflagIIIa particles representing defective virions. However, the IIIa protein contained in mature virions does not contain the N-terminal 6-His tag. The most likely explanation of this finding is that the 6-His tag is cleaved by the adenoviral protease [13] during Ad5GLflagIIIa particle maturation, since (M/I/L)XGX-G and (M/I/L)XGG-X consensus sites[26] are present in the flexible linker between 6-His and FLAG sequences in the engineered N-terminal IIIa extension.

The only part of the 31-residue extension that was unequivocally present in Ad5GLflagIIIa virions was the FLAG peptide. Since there are no detectable epitopes in the flexible linker, it is not possible to determine the exact extension length(s) of IIIa in these virions. The difference peak corresponding to the IIIa N-terminal extension in the 3DEM difference map is just large enough to accommodate a structured FLAG peptide. This would agree with either the FLAG tag

being the only part of the extension present in most of the viral particles, or with the FLAG being the only extension element ordered enough to appear in an averaged, icosahedrally symmetric 3D map. This would be expected, as the FLAG tag is the region of the heterologous peptide closest to the protein native N-terminus.

The positions of adenovirus minor coat proteins in the capsid were established in one of the first combined X-ray/3DEM imaging studies [11]. Recently, a new model has been proposed based on an improved 3DEM map (6 Å *vs.* 25 Å resolution) and secondary structure predictions [20]. Experimental confirmation on the location of as many capsid components as possible will facilitate interpretation of any future adenovirus structural data. Our difference map places the IIIa N-terminal extension close to the newly predicted position for IIIa, at the inner interface between penton base and the peripentonal hexons, and therefore supports the new model. Previous evidence indicated that at least part of IIIa would be exposed on the capsid surface [15]. Both the model by Saban *et al.* [20] and the results reported here refer only to the position of the N-terminal region of polypeptide IIIa, so there might still be a question of whether the C-terminal region of IIIa is actually exposed. Several lines of evidence support the hypothesis that it would also be internally located, possibly in contact with the viral core. First, 15 residues at the C-terminus of pIIIa are cleaved by the adenoviral protease upon virion maturation [13]. The protease is packaged together with the viral genome and uses DNA as a cofactor [27], so presumably the cleavage region has to be close to the core. Second, previous immunoEM assays using a polyclonal serum against IIIa showed labeling of broken particles only [18]. This indicates that not only the N-terminal region, but epitopes throughout the polypeptide chain are not accessible to antibodies from the outside.

Our results show that there is little tolerance for incorporation of extensions at the N-terminus of IIIa, such that only the shortest and less structured ones could be rescued. This implies a relevant role for this region of the protein during assembly. Although heat treatment assays indicated no detectable difference in infectivity between viruses containing wildtype or modified IIIa, it is interesting that the penton density in our Ad5GLflagIIIa 3DEM map is weaker than that of hexons, indicating that in these viruses there is a tendency to release vertex proteins. This is not an uncommon occurrence in adenovirus preparations, depending on the particular purification protocol or storage conditions [17; 19], and so it must be interpreted with caution. However, it is tempting to relate partial penton loss with the modifications introduced in the IIIa N-terminus, which we have mapped to the interface between penton base and the peripentonal hexons on the inner capsid surface.

It has been reported that the Ad2 *ts112* mutant, with just three amino acid changes in polypeptide IIIa, accumulates light particles with little or no DNA when produced at the non-permissive temperature [14; 28]. This is a sign that IIIa has a role in correct genome packaging. Like Ad5GLflagIIIa, Ad2 *ts112* has a tendency to lose pentons, as shown by gaps in the capsid in negatively stained EM images [14]. Structural studies have revealed a striking similarity between adenovirus and the bacterial virus PRD1 [29; 30; 31]. In the PRD1 crystal structure, the position under the vertex at which we find the IIIa N-terminus is occupied by the small (117 residues) capsid protein P16 [32]. P16 links the icosahedral PRD1 capsid to its internal membrane, a structural feature that does not exist in adenovirus. Nevertheless, like we show here for the IIIa N-terminus, five copies of P16 are present at each vertex. They interact with the PRD1 penton base P31, the peripentonal P3 capsomers, and the small cementing protein beneath the capsid edges P30. PRD1 mutants lacking P16 can assemble and package the viral genome with as much

efficiency as the wild type virus, but upon purification they lose the DNA and the vertex structures, much as seen for the Ad2 *ts112* mutant [33]. It seems therefore likely that IIIa is the adenovirus counterpart of PRD1 P16. Both proteins would be required for stabilization of the vertex and the packaged genome, and may be involved in coordination of DNA release events during infection.

Materials and Methods

Construction of recombinant IIIa-modified Ad genomes. To achieve the genetic modifications of the N-terminus of IIIa we cloned a *Pml*I-*Hinc*II DNA fragment (2820 bp) including the IIIa open reading frame (ORF) from plasmid pTG3602 [34] containing the Ad5 genome between *Sma*I and *Hinc*II sites in pBluescriptIISK plasmid (Stratagene, La Jolla, CA) generating plasmid pBsIIIa. To introduce a unique restriction site into the 5'-end of the IIIa gene, PCR primers pIIIaN6His.U (5'-cac cat cac cat cac cat ATG CAA GAC GCA AC) and pIIIaN6His.L (5'-atg gtg atg gtg atg gtg CAT CTG ATC AGA AAC ATC) were designed to encode six consecutive histidines (6-His tag, low case letters) followed by a *Nde*I recognition site (underlined). The pBsIIIa plasmid DNA was used as a template to amplify two DNA fragments, 740-bp and 261-bp long, using the following pairs of primers: pIIIaN.F (5'-CGC GAG GAG GTG GCT ATA GAC TGA) and pIIIaN6His.L, pIIIaN6His.U and pIIIaN.R (5'-TTC GGC CAG CGC GTT TAC GAT C), respectively. DNA fragments were purified and then joined by a second PCR using pIIIaN.F and pIIIaN.R primers. The resultant PCR fragment (983 bp) was digested with *Bsm*I and *Mlu*I and the 738-bp DNA fragment was ligated with the *Bsm*I-*Mlu*I fragment of pBsIIIa plasmid generating plasmid pBs6HIIIa. This plasmid encodes IIIa with both a 6-His tag and *Nde*I site between the first two methionine residues, and formed the basis of all subsequent IIIa modifications.

To incorporate a FLAG octapeptide (Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys) downstream of the 6-His tag in the N-terminus of IIIa, 5'-phosphorylated oligonucleotides AgeIFLAG-F (5'-T **ACC GGT** GAT TAT AAG GAT GAC GAT GAC AAG AGC GC) and AgeIFLAG-R (5'-T AGC GCT CTT GTC ATC GTC ATC CTT ATA ATC **ACC GG**) were annealed to form a DNA duplex encoding a FLAG peptide with an *Age*I site upstream of the FLAG (in bold) and an

*Afe*I site downstream (underlined). This oligo contains *Nde*I-compatible 5'-cohesive ends, and was ligated with *Nde*I-digested pBs6HIIIa to generate shuttle plasmid pBs6HisAgeIFLAGpIIIa, which encodes an 18aa N-terminal IIIa extension. To insert a 45 Å helical spacer [22] between the 6-His tag and FLAG sequence, plasmid pBs6HisAgeIFLAGpIIIa was digested with *Age*I, cutting the plasmid between the 6-His and FLAG sequences. The 45 Å linker was PCR amplified with a forward primer containing an *Age*I restriction site (in bold) (5'-AGCTAA**ACCGGT**GAGACGCGG) and a reverse primer containing an *Xma*I site (in bold) (5'-GCTCAG**CCCGGG**TGGGATCT). A cDNA clone of APOE4 vector from human skin (GenBank I.D.: BC072022, BU839307; I.M.A.G.E. Clone ID:6263100) in pOTB7 vector from ATCC (#8291716) was used as a PCR template to amplify the 45 Å spacer sequence. The spacer PCR product was then digested with *Age*I/*Xma*I to create *Age*I cohesive ends, and ligated into *Age*I-digested pBs6HisAgeIFLAGpIIIa, resulting in shuttle plasmid pBs6HisAgeIFLAGpIIIa45A, which encodes a 60aa extension. To incorporate eGFP into the IIIa N-terminus, the eGFP ORF was PCR-amplified with forward primer containing and *Age*I site (in bold) (5'-CTAGT**ACCGGT**ATGGTGAGCAAGG) and a reverse primer containing an *Xma*I site (in bold) (5'-CTGGA**CCCGGG**CTTGTACAGCTCGT), digested with *Age*I/*Xma*I to create *Age*I cohesive ends, and ligated into *Age*I-digested pBs6HisAgeIFLAGpIIIa to create shuttle plasmid pBs6HisAgeIFLAGpIIIaeGFP, which encodes a 256aa extension. To incorporate a flexible linker and FLAG octapeptide downstream of the 6-His tag in the N-terminus of IIIa, oligonucleotides 5'- TAT TGG TGG AGG CGG TTC AGG CGG AGG TGG CTC TGG CGG TGG CGG ATC CGA TTA TAA GGA TGA CGA TGA CAA GGA and 5'- TAT CCT TGT CAT CGT CAT CCT TAT AAT CGG ATC CGC CAC CGC CAG AGC CAC CTC CGC CTG AAC CGC CTC CAC CAA were synthesized. The oligos were annealed to form a DNA duplex

encoding the $(Gly_4\text{-}Ser)_3$ flexible linker followed by the FLAG peptide while containing the *Nde*I-compatible 5'-cohesive ends and ligated with *Nde*I-digested pBs6HIIIa to generate plasmid pBs6HllFlagIIIa.

To generate genomic plasmids, the *Bsm*I-*Hinc*II DNA fragment isolated from pBs6HIIIa (2675 bp) or pBs6HllFlagIIIa (2750 bp) was used for homologous DNA recombination with Ad5 genome plasmids pTG3602 [34] or pAd5GL [35] linearized with *Pme*I within the IIIa ORF, respectively. For IIIa N-terminal modifications of 18aa (6-His and FLAG), 60 (containing also a helical linker) and 256 (encoding eGFP), shuttle vectors pBs6HisAgeIFLAGpIIIa, pBs6HisAgeIFLAGpIIIa45A and pBs6HisAgeIFLAGpIIIaeGFP were digested with *Eco*R1/*Xho*I/*Xmn*I, and the *Eco*R1/*Xho*I fragments containing the modified IIIa ORF were isolated. These fragments were then used for homologous recombination with *Pme*I-linearized genome plasmid pAdEasy1CMVLuc-ΔpIIIa. This is a pAdEasy1-based Ad5 genome that encodes a CMV promoter/firefly luciferase expression cassette in the E1, but contains a deletion of the N-terminal region of IIIa, encoding the C-terminal 833 bp of the IIIa ORF as well as the native *Pme*I site in the IIIa region.

The resultant pAd5H_6IIIa, pAd5GLflagIIIa, pAd5Luc1-HFpIIIa, pAd5Luc1-H45FpIIIa pAd5Luc1-HeGFPFpIIIa genomic plasmids were digested with *Pac*I to release viral genomic DNA and transfected into 293 cells as described elsewhere [36; 37] to rescue the replication competent Ad5H_6IIIa and all other replication-deficient constructs. All IIIa-modified viruses and the previously described Ad5GL [35] control vector were propagated on 293 cells, purified by centrifugation on CsCl gradients, and dialyzed against PBS (8 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl, and 2.7 mM KCl [pH 7.4]) containing 10% glycerol. The titer of physical viral particles (v.p.) was determined by the methods of Maizel *et al.* [38].

Construction of recombinant IX-modified Ad. Ad vectors Ad5EZ-pIX-Flag-10A-6His F3/3 and Ad5EZ-pIX-Flag-45A-6His F5/3 were constructed based on the Ad5 AdEasy backbone with deleted *E1* and *E3* regions. Both viruses were replication defective and had a cytomegalovirus (CMV) promoter-driven firefly luciferase reporter gene (luc) cassette placed in the *E1* region. Each virus carried an 8 aa FLAG sequence at the C-terminus of polypeptide IX fused to a $GSRGS(H)_6$ sequence motif via a spacer. While Ad5EZ-pIX-Flag-10A-6His F3/3 virus had a 10 aa spacer sequence $P(SA)_4P$ between the FLAG and the $GSRGS(H)_6$ motif, Ad5EZ-pIX-Flag-45A-6His F5/3 virus carried a 37-aa long spacer (EETRARLSKELQAAQARLGAD MEDVCGRLVQYRGEVH) with an α-helical structure also known as the "45Å spacer" [22]. Each spacer was connected to the polypeptide IX-FLAG via a two-amino acid *LA* linker encoded by the sequence of the *Nhe*I site, 3'- adjacent to the FLAG sequence. The viruses also possessed modified fiber protein chimeras with the Ad3 serotype fiber knob domains and either Ad5 (Ad5EZ-pIX-Flag-45A-6His F5/3) or Ad3 (Ad5EZ-pIX-Flag-10A-6His F3/3) serotype fiber shaft domains.

A conventional PCR-based approach followed by cloning of the PCR products into the *Nhe*I site of the shuttle plasmid pShlpIX Flag was used to generate its derivatives carrying the IX-FLAG sequence fused to the 6-His containing motif via two different spacers. To amplify the 10 aa spacer with the $GSRGS(H)_6$ motif a shuttle plasmid pNEB.pk.RGS6HSL containing Ad5 fiber gene with $GSRGS(H)_6$ at the C-terminus was used as a PCR template [39] and the following sense and anti sense PCR primers, respectively: 5'-TGA CAA GCT AGC CCC ATC AGC CTC CGC ATC T; 5'-GCG TTA GCT AGC TTA GGA TCG GGT TTA TTA G. The following sense and anti-sense PCR primers were used to amplify the 45Å-spacer sequence with $GSRGS(H)_6$: 5'-TGA CAA GCT AGC CGA GGA GAC GCG GGC ACG GCT GTC CAA GGA G; 5'-ATG

GTG ATG CGA TCC TCT CGA TCC ATG CAC CTC GCC GCG GTA C. Another anti-sense primer 5'-TTT AAA GCT AGC TTA TTA GTG ATG GTG ATG GTG ATG CGA TCC TCT CGA TCC was used in the second step PCR with the same sense primer to extend the product of the first step reaction and add the sequence corresponding to the GSRGS$(H)_6$. *Nhe*I site recognition sequences are underlined. A cDNA clone of APOE4 vector from human skin (GenBank I.D.: BC072022, BU839307; I.M.A.G.E. Clone ID:6263100) in pOTB7 vector from ATCC (#8291716) was used as a PCR template to amplify the 45 Å spacer sequence.

To introduce the above protein IX modifications into the AdEasy-1 backbone, carrying either F5/3 or F3/3 fiber genes, the corresponding pShuttle plasmids were linearized with *Pme*I (between the homology arm sequences) and recombined with AdEasy-1F5/3 or AdEasy-1F3/3 backbone plasmids in BJ5183 strain of *E. coli* as described previously [40]. In order to generate AdEasy-1 backbones with F5/3 or F3/3 fibers we first generated AdEasy-1 with fiber gene replaced by a unique *Swa*I site and then recombined the *Swa*I-linearized fiberless backbone with two shuttle vectors pKAN3.1F5/3 and pKAN3.1F3/3 as described elsewhere [41]. The fiberless AdEasy-1 was constructed by recombination of the commercial AdEasy-1 vector with pZero2ΔE3.6.9, which in turn was derived from the parental pZero2E3.6.9 plasmid by deletion of the E3 region by a two step PCR approach. Briefly, in the first step two pairs of PCR primers were used to generate a couple of PCR fragments, amplifying 160 and 300 bp size fragments of the Ad5 sequence located between and adjacent to the unique sites *Apa*I and *Swa*I in pZero2E3.6.9. The sense primer of the 160 bp (upstream) PCR fragment and the anti-sense primer of the 300 bp (downstream) PCR fragment contained sequences of *Apa*I and *Swa*I, respectively. The other PCR primer pairs contained an artificially generated *Cla*I site. The PCR products from the first step reaction were digested with *Cla*I and ligated together to serve as a

template for PCR amplification with the external sense (*Apa*I) and anti-sense (*Swa*I) primers used in the first step reactions. The 450 bp chimera product was then digested with *Apa*I and *Swa*I and used to replace *Apa*I-*Swa*I restriction fragment in pZero2E3.6.9 thereby deleting E3 sequence. The resulting E3 deletion (Ad5 nts: 28,133-30,817) contains extra 3 nucleotides at each side of the deletion as compared to the one in the commercial AdEasy-1 (deletion of Ad5 nts: 28,130-30,820).

Protein electrophoresis and Western blot. Samples of CsCl-purified Ad5Luc1-HFpIIIa, Ad5GLFlagIIIa, Ad5H_6IIIa, and Ad5GL virions, or lysates of infected cells, were boiled in Laemmli loading buffer and subjected to 4-20% gradient SDS-polyacrylamide gel electrophoresis (SDS-PAGE) to separate viral proteins. To analyze the N-terminally modified IIIa expressed during viral infection in 293 cells, the cell monolayers were inoculated with Ad5Luc1-HFpIIIa, Ad5GLFlagIIIa, Ad5H_6IIIa, or Ad5GL at a multiplicity of infection (MOI) of 100 v.p./cell. Uninfected and infected cells were harvested 24 hours post-infection, solubilized at 1×10^7 cells/mL in lysis buffer (1% NP-40, 20 mM Tris [pH 8.0], 137 mM NaCl, 10% glycerol, 2 mM EDTA, and 50 µl/mL protease inhibitor cocktail (Sigma, St. Louis, MO)) by gentle rocking at 2 - 8° C for 30 minutes, centrifuged at 14,000 × g for 5 minutes, and 20-µl supernatant aliquots were used immediately for SDS-PAGE. For Western blot analysis, proteins resolved via SDS-PAGE were transferred to PVDF membranes and probed with either anti-Flag M2 mAb (Sigma, St. Louis, MO) or Penta-His mAb (Qiagen, Valencia, CA). Bound M2 and Penta-His mAb were detected with a secondary goat anti-mouse antibody conjugated to alkaline phosphatase (Sigma, St. Louis, MO) by developing the membranes using alkaline phosphatase conjugate substrate kit (Bio-Rad Laboratories, Hercules, CA) as recommended by the manufacturer.

Gene transfer assay. Monolayers of 293 cells grown in a 24-well plate (3 - 5 × 10^5 cells/well) were inoculated with 0.2-ml aliquots of Ad vectors that were incubated at 45°C for different time intervals or untreated control virus at MOI of 100 v.p./cell. Ad vectors were allowed to be internalized for 30 min and then, infection medium was aspirated, the cells were washed with PBS, and incubated in growth medium at 37°C to allow expression of luciferase reporter gene. Cells were lysed 20 h postinfection and luciferase activity was analyzed by using the Promega (Madison, Wis.) luciferase assay as we previously described [40].

Immuno-electron microscopy. Samples of Ad5GLflagIIIa and Ad5GL to be used for electron microscopy were stored at -70ºC in small aliquots, to avoid repeated freezing and thawing. Viruses were adsorbed for 5 min onto glow-discharged, Formvar/carbon coated nickel grids. The Ad5GL vector, which is structurally wild type, was used as a negative control. Ad5EZ-pIX-Flag-10A-6His-F3/3 and Ad5EZ-pIX-Flag-45A-6His-F3/3 vectors with 6-His and FLAG-containing extensions of different lengths at the C-terminal end of polypeptide IX, were used as positive controls. Grids were rinsed in TBS (20 mM Tris-HCl pH 7.8, 150 mM NaCl) for 2 min and unspecific binding blocked with TBG (1% cold water fish gelatin, 0.1% bovine serum albumin (BSA) in TBS) for 15 min. Grids were then incubated for 50 min with mAb against either the 6-His tag (Clontech ref. 631212) or the FLAG peptide (Sigma M2), in dilutions ranging from 1:25 to 1:100 in TBS/BSA (100:1). Grids were rinsed three times with TBG and incubated with 15% 10 nm goat anti-mouse IgG-gold conjugate (British Biocell International EM.GAF10) in TBS/BSA for 30 min. After three rinses in TBG, and three in TBS, grids were stained with 2% uranyl acetate and examined in a JEOL 1200 EX-II electron microscope. For gold particle count, micrographs were taken at 25x magnification from areas selected at low magnification.

Cryo-electron microscopy. Ad5GL and A5GLflagIIIa virus samples were dialyzed for 1 hour at 4ºC against 10 mM phosphate buffer pH 7.2, 150 mM NaCl, applied to freshly carbon-coated, glow discharged Quantifoil R2/4 300 mesh Cu/Rh grids, and vitrified in liquid ethane using a Leica CPC plunger. Grids were mounted in a Gatan 626 cryostage and examined in a FEI Tecnai G2 FEG microscope operating at 200 kV. Micrographs were recorded on Kodak SO-163 film under low dose conditions at a nominal magnification of 50,000x, and digitized in a Zeiss Photoscan TD scanner using a step size of 7 μm (1.4 Å in the sample).

Three-dimensional reconstruction and difference mapping. Image processing and three-dimensional reconstruction were performed using the software packages Xmipp [42; 43] and SPIDER [44]. The contrast transfer function (CTF) parameters of each micrograph were determined from its rotationally averaged power spectrum, obtained by patch averaging. Micrographs free of drift and astigmatism were selected and downsampled to a final pixel size of 4.2 Å/px. Only those micrographs within a common underfocus range (1.7 to 5 μm) for both Ad5GL and Ad5GLflagIIIa were further used for calculation of the 3D maps (66 micrographs for Ad5GL, 89 for Ad5GLflagIIIa).

Initial sets of 3120 (Ad5GL) and 3870 (Ad5GLflagIIIa) particles were manually selected, extracted into 275x275 pixel boxes, and normalized. Images were sorted into defocus groups and iterative projection matching was carried out within each group, using as a model the 3D map calculated in the previous iteration, affected by the corresponding CTF [45]. For every iteration, separate 3D reconstructions were calculated in each defocus group and then combined into a Wiener filter CTF corrected map [46]. The initial model for the first projection matching iteration was a previously obtained Ad5 3D map [18]. Particles giving the worst correlation coefficient values during projection matching were rejected. A total of 2575 (Ad5GL) and 3171

(Ad5GLflagIIIa) particles were included in the final 3D reconstructions. The resolution of the 3D maps was estimated using the FSC=0.5 criterion, considering only a shell loosely containing the icosahedrally ordered part of the capsid (radii 294 to 504 Å). Icosahedral symmetry was imposed throughout projection matching and reconstruction.

Single particle reconstruction (SPR) based classification. A 3D map was calculated from every single image in the Ad5GLflagIIIa dataset filtered at 30 Å, using the orientations determined by projection matching, and with icosahedral symmetry enforced. We have previously shown that the use of the iterative reconstruction algorithm ART [47] with a low-resolution starting volume considerable reduces artifacts in SPRs, a particularly critical feature when the region of interest for classification falls close to the icosahedral symmetry axes [18]. Therefore, we calculated SPRs with ART using a starting volume created from a dataset containing two images from each defocus group (44 images in total), filtered to 30 Å resolution. SPR calculations were carried out in an HP Proliant cluster with 32 nodes and 2 Intel Xeon 3.06 GHz CPUs per node. The average density value of each SPR within a mask covering the penton was calculated and used to select two SPR subsets, each containing approximately 1/3rd of the complete dataset, with either the strongest or the weakest penton density values. The corresponding experimental images were then used to calculate two 3D maps, one with the highest penton density images and another one with the lowest (977 and 972 images respectively). For the Ad5GL control, 980 experimental images were randomly selected covering a defocus range similar to that of the SPR-selected Ad5GLflagIIIa subsets, and a 3D reconstruction calculated. Penton occupancy for each map was defined as the penton/hexon density ratio, with the value for Ad5GL taken as 100%. The ratios were calculated for all maps

by comparing their average density values within masks that segmented either the penton or one of the peripentonal hexons.

Difference maps. To calculate the various difference maps, 3D maps to be compared were filtered to a common resolution limit, grayscale normalized within radii 294 to 504 Å, centered, and brought to a common scale before subtraction. The difference values were normalized to average 0 and standard deviation 1 within the same radii.

Fitting of high resolution structures. A FLAG peptide structure model was taken from the crystal structure of peptidylglycan hydrolase ALE-1 (PDB ID 1R77) and fitted to the Ad5GLflagIIIa-Ad5GL difference peak, using UCSF Chimera [48]. The structures of penton base and the peripentonal hexons from a previously reported Ad5 quasi-atomic model [17] (PDB ID 2BLD) were obtained from the VIPERDB virus structure database [49]. After manually placing the models onto the Ad5GL 3D map, the fit was refined using UCSF Chimera. Surfaces for the high resolution structures were calculated and displayed with the UCSF Chimera Multiscale Models tool.

3DEM database submission and entry IDs. The Ad5GL and Ad5GLflagIIIa 3DEM maps have been deposited at the Macromolecular Structure Database (MSD, http://www.ebi.ac.uk/msd) with accession codes EMD-1489 and EMD-1490, respectively.

Acknowledgements

This work was supported by grants from the Ministerio de Educación y Ciencia of Spain (BFU2007-60228 to C. S. M. and BIO2007-67150-C03-03 to R. M.); the Spanish Fondo de Investigaciones Sanitarias (04/0863 to R. M.); the EU 3DEM Network of Excellence (FP6-502828); and NIH (5R01CA111569 and 5T32CA075930 to D. T. C. and R21CA116525 to I. P. D).

We gratefully acknowledge J.M. Carazo (CNB-CSIC) for generous support during initial stages of the project. We are also grateful to J. Martín-Benito (CNB-CSIC) for sharing computational equipment, and the Computer Architecture Department of the Almeria University for providing access to the HP Proliant cluster.

References

1. Benkö, M. & Harrach, B. (2003). Molecular evolution of adenoviruses. In *Adenoviruses: Model and Vectors in Virus Host Interactions. Current Topics in Microbiology and Immunology* (Doerfler, W. & Böhm, P., eds.), Vol. 272, pp. 3-36. Springer-Verlag, Heidelberg.
2. Kinchington, P. R., Romanowski, E. G. & Jerold Gordon, Y. (2005). Prospects for adenovirus antivirals. *J. Antimicrob. Chemother.* **55**, 424-429.
3. Dobbelstein, M. (2004). Replicating adenoviruses in cancer therapy. *Curr Top Microbiol Immunol* **273**, 291-334.
4. Shiver, J. W., Fu, T.-M., Chen, L., Casimiro, D. R., Davies, M.-E., Evans, R. K., Zhang, Z.-Q., Simon, A. J., Trigona, W. L., Dubey, S. A., Huang, L., Harris, V. A., Long, R. S., Liang, X., Handt, L., Schleif, W. A., Zhu, L., Freed, D. C., Persaud, N. V., Guan, L., Punt, K. S., Tang, A., Chen, M., Wilson, K. A., Collins, K. B., Heidecker, G. J., Fernandez, V. R., Perry, H. C., Joyce, J. G., Grimm, K. M., Cook, J. C., Keller, P. M., Kresock, D. S., Mach, H., Troutman, R. D., Isopi, L. A., Williams, D. M., Xu, Z., Bohannon, K. E., Volkin, D. B., Montefiori, D. C., Miura, A., Krivulka, G. R., Lifton, M. A., Kuroda, M. J., Schmitz, J. E., Letvin, N. L., Caulfield, M. J., Bett, A. J., Youil, R., Kaslow, D. C. & Emini, E. A. (2002). Replication-incompetent adenoviral vaccine vector elicits effective anti-immunodeficiency-virus immunity. *Nature* **415**, 331-5.
5. van Oostrum, J. & Burnett, R. M. (1985). Molecular Composition of the Adenovirus Type-2 Virion. *Journal of Virology* **56**, 439-448.

6. Rux, J. J. & Burnett, R. M. (2000). Type-specific epitope locations revealed by X-ray crystallographic study of adenovirus type 5 hexon. *Molecular Therapy* **1**, 18-30.

7. Zubieta, C., Schoehn, G., Chroboczek, J. & Cusack, S. (2005). The structure of the human adenovirus 2 penton. *Mol Cell* **17**, 121-35.

8. Xia, D., Henry, L. J., Gerard, R. D. & Deisenhofer, J. (1994). Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution. *Structure* **2**, 1259-70.

9. van Raaij, M. J., Louis, N., Chroboczek, J. & Cusack, S. (1999). Structure of the human adenovirus serotype 2 fiber head domain at 1.5 Å resolution. *Virology* **262**, 333-43.

10. Burnett, R. M. (1985). The structure of the adenovirus capsid. II. The packing symmetry of hexon and its implications for viral architecture. *J Mol Biol* **185**, 125-43.

11. Stewart, P. L., Fuller, S. D. & Burnett, R. M. (1993). Difference Imaging of Adenovirus - Bridging the Resolution Gap between X-Ray Crystallography and Electron-Microscopy. *EMBO Journal* **12**, 2589-2599.

12. Boudin, M. L., D'Halluin, J. C., Cousin, C. & Boulanger, P. (1980). Human adenovirus type 2 protein IIIa. II. Maturation and encapsidation. *Virology* **101**, 144-56.

13. Webster, A., Russell, S., Talbot, P., Russell, W. C. & Kemp, G. D. (1989). Characterization of the adenovirus proteinase: substrate specificity. *J Gen Virol* **70 (Pt 12)**, 3225-34.

14. D'Halluin, J. C., Milleville, M., Boulanger, P. A. & Martin, G. R. (1978). Temperature-sensitive mutant of adenovirus type 2 blocked in virion assembly: accumulation of light intermediate particles. *J Virol* **26**, 344-56.

15. Everitt, E., Lutter, L. & Philipson, L. (1975). Structural proteins of adenoviruses. XII. Location and neighbor relationship among proteins of adenovirion type 2 as revealed by enzymatic iodination, immunoprecipitation and chemical cross-linking. *Virology* **67**, 197-208.

16. Newcomb, W. W. & Brown, J. C. (1988). Use of Ar+ plasma etching to localize structural proteins in viruses: studies with adenovirus 2. *Anal Biochem* **169**, 279-86.

17. Fabry, C. M., Rosa-Calatrava, M., Conway, J. F., Zubieta, C., Cusack, S., Ruigrok, R. W. & Schoehn, G. (2005). A quasi-atomic model of human adenovirus type 5 capsid. *Embo J* **24**, 1645-54.

18. Scheres, S. H. W., Marabini, R., Lanzavecchia, S., Cantele, F., Rutten, T., Fuller, S. D., Carazo, J. M., Burnett, R. M. & San Martín, C. (2005). Classification of single-projection reconstructions for cryo-electron microscopy data of icosahedral viruses. *J Struct Biol* **151**, 79-91.

19. Marsh, M. P., Campos, S. K., Baker, M. L., Chen, C. Y., Chiu, W. & Barry, M. A. (2006). CryoEM of Protein IX-Modified Adenoviruses Suggests a New Position for the C-terminus of Protein IX. *J Virol* **80**, 11881-11886.

20. Saban, S. D., Silvestry, M., Nemerow, G. R. & Stewart, P. L. (2006). Visualization of α-helices in a 6 Å resolution cryoEM structure of adenovirus allows refinement of capsid protein assignments. *J Virol* **80**, 12049-12059.

21. Schoehn, G., El Bakkouri, M., Fabry, C. M., Billet, O., Estrozi, L. F., Le, L., Curiel, D. T., Kajava, A. V., Ruigrok, R. W. & Kremer, E. J. (2008). Three-dimensional structure of canine adenovirus serotype 2 capsid. *J Virol* **82**, 3192-203.

22. Vellinga, J., Rabelink, M. J., Cramer, S. J., van den Wollenberg, D. J., Van der Meulen, H., Leppard, K. N., Fallaux, F. J. & Hoeben, R. C. (2004). Spacers increase the accessibility of peptide ligands linked to the carboxyl terminus of adenovirus minor capsid protein IX. *J Virol* **78**, 3470-9.

23. Le, L. P., Everts, M., Dmitriev, I. P., Davydova, J. G., Yamamoto, M. & Curiel, D. T. (2004). Fluorescently labeled adenovirus with pIX-EGFP for vector detection. *Mol Imaging* **3**, 105-16.

24. Le, L. P., Le, H. N., Nelson, A. R., Matthews, D. A., Yamamoto, M. & Curiel, D. T. (2006). Core labeling of adenovirus with EGFP. *Virology* **351**, 291-302.

25. Conway, J. F., Cheng, N., Zlotnick, A., Stahl, S. J., Wingfield, P. T. & Steven, A. C. (1998). Localization of the N terminus of hepatitis B virus capsid protein by peptide-based difference mapping from cryoelectron microscopy. *Proc Natl Acad Sci U S A* **95**, 14622-7.

26. Diouri, M., Keyvani-Amineh, H., Geoghegan, K. F. & Weber, J. M. (1996). Cleavage efficiency by adenovirus protease is site-dependent. *J Biol Chem* **271**, 32511-4.

27. Weber, J. M. (1999). Role of endoprotease in adenovirus infection. In *Adenoviruses : basic biology to gene therapy* (Seth, P., ed.), pp. 79-83. R.G. Landes, Austin, Tex., U.S.A.

28. Chroboczek, J., Viard, F. & D'Halluin, J. C. (1986). Human adenovirus 2 temperature-sensitive mutant 112 contains three mutations in the protein IIIa gene. *Gene* **49**, 157-60.

29. Benson, S. D., Bamford, J. K. H., Bamford, D. H. & Burnett, R. M. (1999). Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. *Cell* **98**, 825-833.

30. Butcher, S. J., Bamford, D. H. & Fuller, S. D. (1995). DNA packaging orders the membrane of bacteriophage PRD1. *EMBO Journal* **14**, 6078-86.

31. San Martín, C., Huiskonen, J. T., Bamford, J. K., Butcher, S. J., Fuller, S. D., Bamford, D. H. & Burnett, R. M. (2002). Minor proteins, mobile arms and membrane-capsid interactions in the bacteriophage PRD1 capsid. *Nat Struct Biol* **9**, 756-63.

32. Abrescia, N. G., Cockburn, J. J., Grimes, J. M., Sutton, G. C., Diprose, J. M., Butcher, S. J., Fuller, S. D., San Martín, C., Burnett, R. M., Stuart, D. I., Bamford, D. H. & Bamford, J. K. (2004). Insights into assembly from structural analysis of bacteriophage PRD1. *Nature* **432**, 68-74.

33. Jaatinen, S. T., Viitanen, S. J., Bamford, D. H. & Bamford, J. K. (2004). Integral membrane protein P16 of bacteriophage PRD1 stabilizes the adsorption vertex structure. *J Virol* **78**, 9790-7.

34. Chartier, C., Degryse, E., Gantzer, M., Dieterle, A., Pavirani, A. & Mehtali, M. (1996). Efficient generation of recombinant adenovirus vectors by homologous recombination in Escherichia coli. *J Virol* **70**, 4805-10.

35. Seki, T., Dmitriev, I., Kashentseva, E., Takayama, K., Rots, M., Suzuki, K. & Curiel, D. T. (2002). Artificial extension of the adenovirus fiber shaft inhibits infectivity in coxsackievirus and adenovirus receptor-positive cell lines. *J Virol* **76**, 1100-8.

36. Dmitriev, I., Krasnykh, V., Miller, C. R., Wang, M., Kashentseva, E., Mikheeva, G., Belousova, N. & Curiel, D. T. (1998). An adenovirus vector with genetically modified fibers demonstrates expanded tropism via utilization of a coxsackievirus and adenovirus receptor-independent cell entry mechanism. *J Virol* **72**, 9706-13.

37. Krasnykh, V., Dmitriev, I., Mikheeva, G., Miller, C. R., Belousova, N. & Curiel, D. T. (1998). Characterization of an adenovirus vector containing a heterologous peptide epitope in the HI loop of the fiber knob. *J Virol* **72**, 1844-52.

38. Maizel, J. V., Jr., White, D. O. & Scharff, M. D. (1968). The polypeptides of adenovirus. I. Evidence for multiple protein components in the virion and a comparison of types 2, 7A, and 12. *Virology* **36**, 115-25.

39. Douglas, J. T., Miller, C. R., Kim, M., Dmitriev, I., Mikheeva, G., Krasnykh, V. & Curiel, D. T. (1999). A system for the propagation of adenoviral vectors with genetically modified receptor specificities. *Nat Biotechnol* **17**, 470-5.

40. Dmitriev, I. P., Kashentseva, E. A. & Curiel, D. T. (2002). Engineering of adenovirus vectors containing heterologous peptide sequences in the C terminus of capsid protein IX. *J Virol* **76**, 6893-9.

41. Krasnykh, V., Belousova, N., Korokhov, N., Mikheeva, G. & Curiel, D. T. (2001). Genetic targeting of an adenovirus vector via replacement of the fiber protein with the phage T4 fibritin. *J Virol* **75**, 4176-83.

42. Marabini, R., Masegosa, I. M., San Martín, M. C., Marco, S., Fernández, J. J., de la Fraga, L. G., Vaquerizo, C. & Carazo, J. M. (1996). Xmipp: An Image Processing Package for Electron Microscopy. *J Struct Biol* **116**, 237-40.

43. Sorzano, C. O., Marabini, R., Velázquez-Muriel, J., Bilbao-Castro, J. R., Scheres, S. H., Carazo, J. M. & Pascual-Montano, A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. *J Struct Biol* **148**, 194-204.

44. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. & Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *Journal of Structural Biology* **116**, 190-9.

45. Penczek, P. A., Grassucci, R. A. & Frank, J. (1994). The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy* **53**, 251-70.

46. Penczek, P. A., Zhu, J., Schröeder, R. & Frank, J. (1997). Three Dimensional Reconstruction with Contrast Transfer Compensation from Defocus Series. *Special Issue on Signal and Image Processing, Scanning Microscopy* **11**, 147-154.

47. Marabini, R., Herman, G. T. & Carazo, J. M. (1998). 3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs). *Ultramicroscopy* **72**, 53-65.

48. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12.

49. Reddy, V. S., Natarajan, P., Okerberg, B., Li, K., Damodaran, K. V., Morton, R. T., Brooks, C. L., 3rd & Johnson, J. E. (2001). Virus Particle Explorer (VIPER), a website for virus capsid structures and their computational analyses. *J Virol* **75**, 11943-7.

Figure Legends

Fig 1. Schematic representation of modified IIIa proteins indicating non-native N-terminal extensions size and structure.

Fig 2. (A) Detection of 6-His- or FLAG-tagged IIIa proteins expressed by the generated Ad variants. Samples of 293 cells infected with Ad5GL (lane 2), Ad5H_6IIIa (lane 3), Ad5Luc1-HFpIIIa (lane 4), or Ad5GLflagIIIa (lane 5) were analyzed by Western blot using anti-6-His Penta-His (upper panel) or anti-Flag M2 mAb (lower panel). Molecular masses of marker proteins (lane 1) indicated in kilodaltons on the left. **(B)** Detection of modified protein IIIa in the Ad capsid. Samples of defective particles of Ad5GL (lane 2), Ad5H_6IIIa (lane 3), Ad5Luc1-HFpIIIa (lane 4), and Ad5GLflagIIIa (lane 5) or mature Ad5GL (lane 6), Ad5H_6IIIa (lane 7), Ad5Luc1-HFpIIIa (lane 8), and Ad5GLflagIIIa (lane 9) virions purified by centrifugation on CsCl gradient were analyzed by Western blot using anti-6-His Penta-His (upper panel) or anti-Flag M2 mAb (lower panel). Molecular masses of marker proteins (lane 1) indicated in kilodaltons on the left.

Fig. 3. Thermostability of IIIa-modified Ad vectors. Aliquots of Ad5Luc1-HFpIIIa, Ad5GLflagIIIa, and Ad5GL vectors were incubated at 45°C for different time intervals and then used to infect 293 cells. Results are presented as the percentages of luciferase activity detected in the cells infected with a heat-treated viral sample with respect to luciferase activity determined in the cells infected with untreated virus (100%). Each bar represents the cumulative mean ± SD of triplicate determinations. Some error bars depicting standard deviations are smaller than the symbols.

Fig 4. Immuno-EM of Ad5GL, Ad5GLflagIIIa, and Ad5EZ-pIX-Flag-45A-6His F5/3 labeled with anti-FLAG mAb diluted 1:25. Arrows indicate gold particles in the Ad5GL and Ad5GLflagIIIa panels, both practically devoid of label. The bar represents 200 nm.

Fig 5: (A) Central sections of the Ad5GL and Ad5GLflagIIIa 3D maps and their difference, as indicated. Highest density is white. Arrows point to the strongest positive difference peak, located underneath the penton. White symbols indicate the icosahedral 5-fold (pentagon) and 2-fold (ovals) symmetry axes. Sections are perpendicular to an icosahedral 2-fold symmetry axis. The scale bar represents 100 Å. **(B)** Fourier shell correlation curves for the Ad5GL and Ad5GLflagIIIa reconstructions. Both maps were filtered to the lowest common resolution (14 Å) before calculating the difference map. **(C)** Central sections of the Ad5GLflagIIIa maps calculated with those particles with highest **(high)** or lowest **(low)** penton density, and their difference with the Ad5GL_subset map (**high-subset** and **low-subset**). Arrows indicate difference peaks at the same position as that in the global difference map (A).

Fig 6: (A) Surface rendering of the Ad5GL 3DEM map contoured at 1σ level (grey) and the Ad5GLflagIIIa – Ad5GL difference map contoured at 3σ (red). The view is looking at the penton (indicated with a pentagon) from inside the capsid. The bar represents 100 Å. One of the 60 difference peaks attributed to the polypeptide IIIa N-terminal extension is indicated with a black circle. **(B)** A zoom in the difference peak indicated with the circle in (A), with the atomic structure of the FLAG peptide (blue sticks) fitted into the density (red mesh). **(C)** as in (A), but showing a transversal section of the capsid across the penton. The penton base (**pb**) and two

peripentonal hexons (**ph**) are labeled. **(D)** and **(E)** show the crystal structures for penton base (yellow) and peripentonal hexons (light blue) represented as surfaces, and fitted into the Ad5GL map, with the difference map in red. The views are as in (A) and (C), respectively. A slab of the Ad5GL map is shown in transparent grey in (E).

Table 1. ImmunoEM assay quantitation

	Number of gold labels per virus particle			
	Anti-6-His		**Anti-FLAG**	
	Intact virions	**Disrupted virions**	**Intact virions**	**Disrupted virions**
Ad5GL	0.04 ± 0.20 (N=100)	0.11 ± 0.33 (N=9)	0.00 ± 0.00 (N=100)	0.21 ± 0.49 (N=53)
Ad5GLflagIIIa	0.10 ± 0.30 (N=100)	0.48 ± 0.94 (N=27)	0.02 ± 0.14 (N=100)	0.38 ± 0.77 (N=40)
Ad5EZ-pIX-Flag-45A-6His-F3/3	6.79 ± 2.30 (N=100)	5.35 ± 1.97 (N=17)	4.92 ± 2.04 (N=100)	3.60 ± 2.51 (N=35)
Ad5EZ-pIX-Flag-10A-6His F3/3	3.12 ± 1.67 (N=100)	3.64 ± 1.69 (N=14)	4.52 ± 1.93 (N=100)	3.77 ± 2.01 (N=26)

Figure

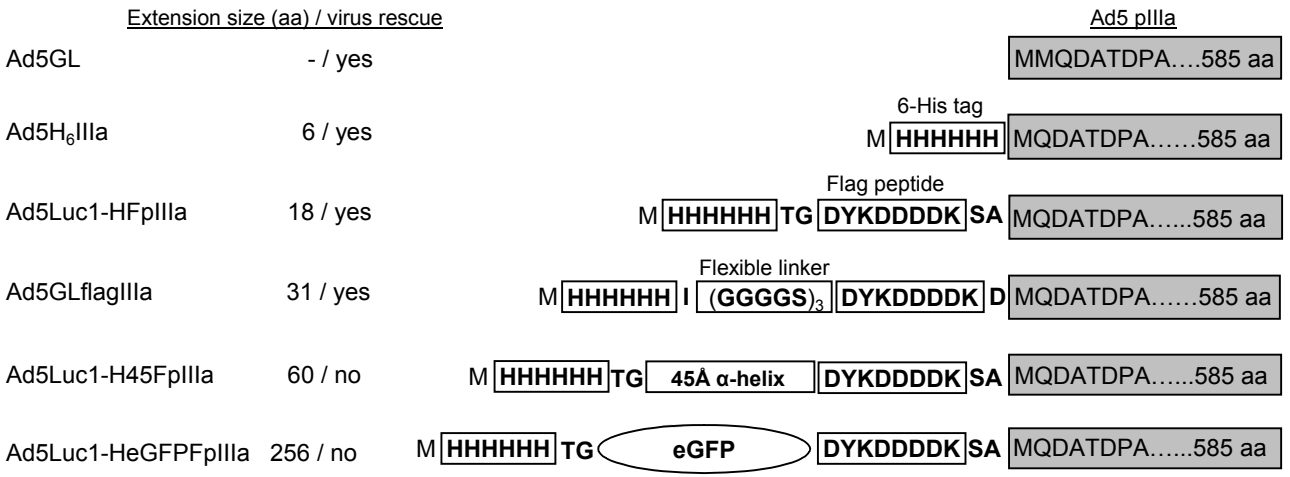

Extension size (aa) / virus rescue
Ad5 pIIIa
Ad5GL
- / yes
MMQDATDPA….585 aa
6-His tag
Ad5H6IIIa
6 / yes
M HHHHHH MQDATDPA……585 aa
Flag peptide
Ad5Luc1-HFpIIIa
18 / yes
M HHHHHH TG DYKDDDDK SA MQDATDPA……585 aa
Flexible linker
Ad5GLflagIIIa
31 / yes
M HHHHHH I (GGGGS)3 DYKDDDDK D MQDATDPA……585 aa
Ad5Luc1-H45FpIIIa
60 / no
M HHHHHH TG 45Å α-helix DYKDDDDK SA MQDATDPA……585 aa
Ad5Luc1-HeGFPFpIIIa
256 / no
M HHHHHH TG eGFP DYKDDDDK SA MQDATDPA……585 aa

Fig. 2

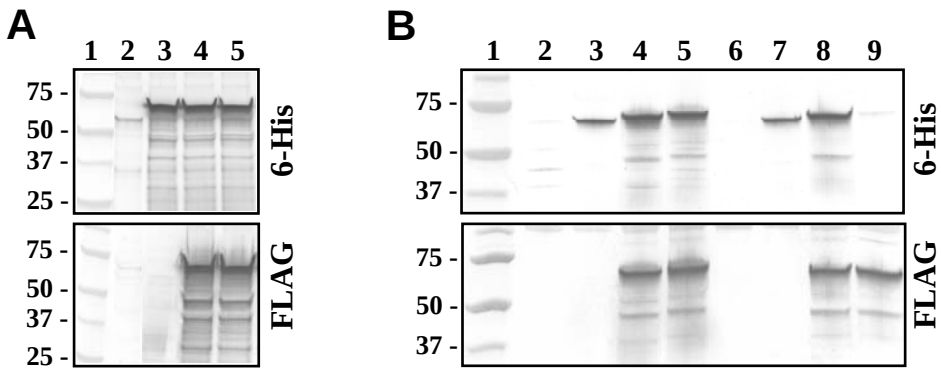
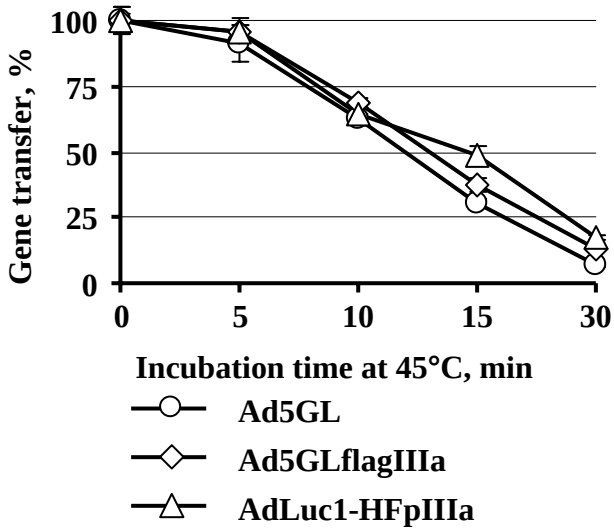
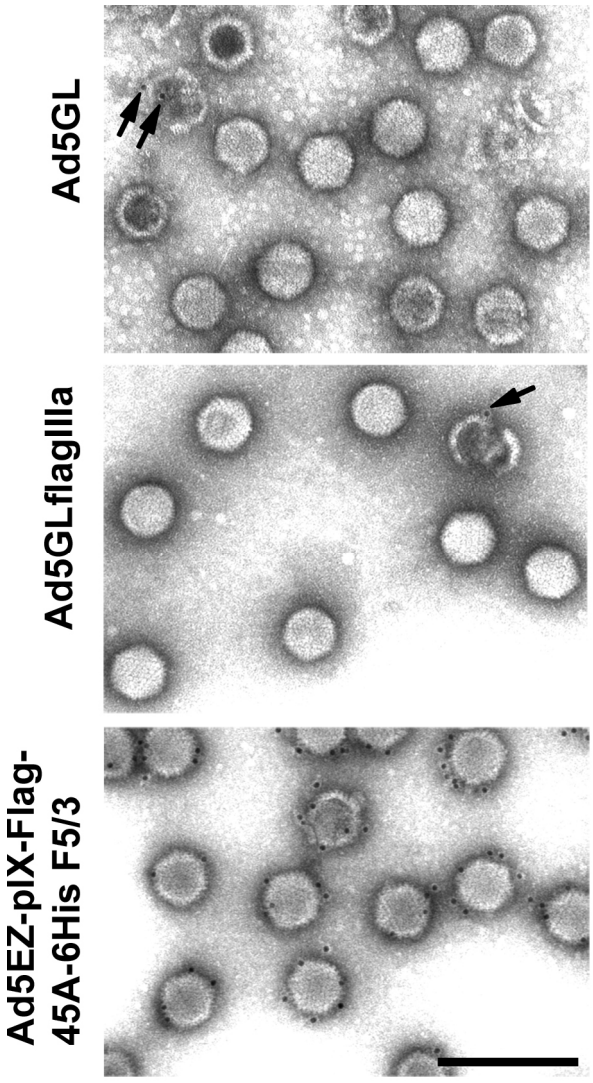

A
1 2 3 4 5
75 -
50 -
37 -
25 -
6-His
75 -
50 -
37 -
25 -
FLAG
B
1 2 3 4 5 6 7 8 9
75 -
50 -
37 -
6-His
75 -
50 -
37 -
FLAG

Fig. 3

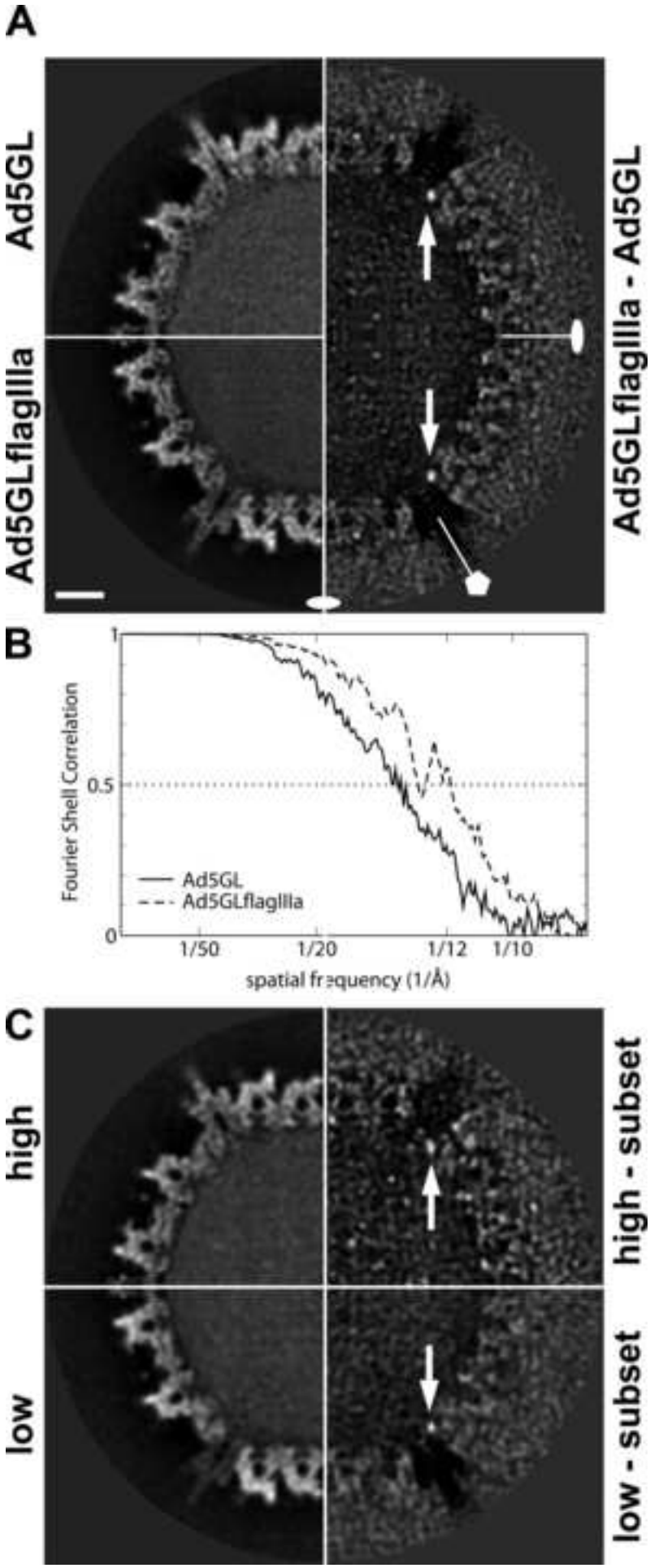

Gene transfer, %
100
75
50
25
0
0
5
10
15
30
Incubation time at 45°C, min
Ad5GL
Ad5GLflagIIIa
AdLuc1-HFpIIIa

Figure



Figure

Click here to download high resolution image



Figure

Click here to download high resolution image