CRYO-EM STRUCTURES OF ADENOVIRUS PROVIDE INSIGHT INTO VECTOR DESIGN, VIRUS ASSEMBLY, AND CAPSID DISASSEMBLY

By

Susan Dora Saban

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Molecular Physiology and Biophysics

December, 2006

Nashville, Tennessee

Approved:

Professor Albert H. Beth

Professor Hassane S. Mchaourab

Professor David W. Piston

Professor Charles R. Sanders, II

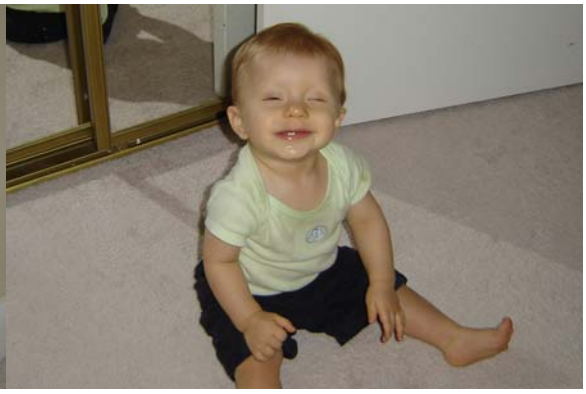
Professor Terrence S. Dermody

For my loving husband, Eloy, and my darling son, Dillan.

You are my inspiration.





ACKNOWLEDGEMENTS

First and foremost I would like to thank my mentor, Phoebe Stewart. I could not have picked a better person under whom to do my thesis research. She is a brilliant scientist whose creative ideas have helped me to navigate obstacles in my research. I admire her buoyant outlook on research and her drive to push on despite the challenges this field can present. I am grateful for the opportunity she gave me of transferring to Vanderbilt University from UCLA during my first year of graduate school. I would not have otherwise had the opportunity to work with such "state-of-the-art" microscopes and computational facilities. I also appreciate the many occasions to travel to scientific meetings and present my research. Lastly, I am very thankful for her flexibility when my life dramatically changed a little less than a year ago. Being able to finish this "chapter" in my life is very satisfying.

I am grateful for the pleasure of working with those I encountered during my thesis research. I appreciate the feedback my thesis committee has given along the way, which has helped me to develop scientifically. It was a true joy working with fellow members of the Stewart Laboratory. Dewight's dedication to his work was especially inspiring when I was working out plunging conditions for adenovirus. He also critically read many documents and asked thoughtful questions at group meetings, which reminded me many times to deliver my message to the general audience. Jian is a savvy computer programmer, and I appreciated his expertise when I was learning to program in C. I am indebted to Mariena for assisting in the collection of my last adenovirus data set. She has been a good friend and confidant. We shared many spiritual conversations that you don't

normally hear in the scientific community. I would also like to thank past members of the Stewart Lab: Yeshi Mikyas, Miriam Makabi, and Steve Chiu for first getting me hooked on cryoEM. Also, thank you to Rebecca Erickson, whose friendship was invaluable over the course of graduate school. Lastly, my thesis research and completion of this dissertation would not have been possible without the love and support of my family. I am especially grateful to my mother who stepped in for over a month to allow me to focus on writing and defending.

I am grateful for financial support from the Molecular Biophysics Training Grant during the second and third years of graduate school. Being a part of the Molecular Biophysics Training Program seminar series and courses was an enlightening experience. I would like to thank Lance Gritton for preparation of the Ad35F vector; the NRAMM staff for their assistance with Leginon and for the negatively-stain TMV calibration grid; Niko Grigorieff for assistance with Frealign; and Jarrod Smith and the ACCRE staff for computer support.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Adenoviruses (Ads) are a large family of non-enveloped, double-stranded DNA viruses that are a significant cause of acute respiratory, gastrointestinal, and ocular infections in human hosts. Adenoviruses can present more serious illnesses in immunocompromised individuals with AIDS or those undergoing chemotherapy or bone marrow transplants. Adenovirus has served as an important system for uncovering a wide array of cellular and molecular processes including RNA splicing and cell cycle regulation. Adenoviruses currently represent a significant proportion of the viral vectors used for gene therapy and vector-based vaccine applications due to high titer production, wide tropism, ability to infect dividing and non-dividing host cells, high stability *in vivo*, absence of chromosomal integration, and high infection efficiency (118).

Adenovirus-based Vectors

Adenovirus vectors have been utilized as therapeutic agents in a host of diseases including cancer. To address the genetic lesions that lead to cancer, numerous adenovirus-vector based gene therapy approaches have been developed with the newest being oncolytic therapy or virotherapy. Virotherapy relies on the ability of the virus to replicate in and lyse tumor cells. Adenoviruses have a lytic life cycle, but early vectors were replication-defective meaning that infection of the target host cell was a terminal event. These vectors would have low efficacy in virotherapy where infection of

surrounding tumor cells is essential. Conditionally replicating Adenovirus vectors (CRAds) with mutations in the E1B gene were created to allow vectors to replicate in p53 deficient tumors (6). The E1B 55-kDa protein binds and inactivates p53 to block apoptosis long enough to allow viral replication. In the absence of the E1B gene in CRAds, healthy cells are thought to undergo apoptosis when infected, whereas tumor cells would allow replication, lysis, and infection of the surrounding tumor cells. First generation CRAds are well tolerated by patients with various cancers in phase I and II clinical trials (51, 70). Unfortunately, the efficacy of CRAd virotherapy is limited by the variable expression of the attachment receptor on tumor cells (58, 106, 151). Also, since the attachment receptor is abundant on liver cells, particle uptake and sequestration in the liver are hepatotoxic and activate the host immune response (94).

Vector Retargeting

Efficient Adenovirus cell entry involves association with two receptors. Most of the 51 different Adenovirus serotypes, including Ad5 used in CRAds, recognize CAR (Coxsackievirus and Adenovirus Receptor) for cell attachment and subsequently α_v integrins for internalization (122). The immune response resulting from vector uptake and sequestration in the liver may be circumvented by redirecting the vector away from liver cells and towards tumor cells. Therefore, to bypass the liver, the main strategy is to block the Ad5 fiber from attaching to liver cells via CAR. To achieve this end three approaches have been undertaken: first, development of bispecific retargeting ligands such as bispecific fusion antibodies that recognize the adenovirus fiber and tumor cell markers (75); second, vectors with genetically modified capsid proteins (e.g. protein IX) which

can incorporate biotinylated cellular targeting ligands (14); and third, pseudotyped vectors.

Pseudotyped vectors take advantage of the fact that some adenovirus serotypes recognize attachment receptors other than CAR. Subgroup B and D adenoviruses, including Ad35 (subgroup B) utilize an alternate attachment receptor, CD46/MCP (membrane cofactor protein) (41). CD46 is a member of the family of complement regulatory proteins and an attachment receptor for several microbial pathogens as well as for a subset of adenovirus serotypes. In humans, CD46 is expressed at low levels in all nucleated cells, but it is overexpressed in malignant tumors (38). Ad35 is a rare human pathogen that is associated with infections in immunocompromised patients. Ad35 has received attention as a potential gene delivery vector since it can overcome pre-existing anti-adenovirus immunity, and it can infect non-CAR expressing cells such as hematopoietic cells (87, 114). In a transgenic mouse model with CD46 expressed at levels similar to humans, an adenovirus vector with the capsid of Ad5 and the fiber of Ad35 (Ad35F) preferentially targets tumors overexpressing CD46 and has anti-tumor properties (83).

Adenovirus Structure

A detailed structural understanding of adenovirus would complement vector design; however, a high resolution structural analysis of intact adenovirus has been confounded by the large size and complexity of the virus. Adenoviruses are composed of eleven different structural proteins and a double-stranded DNA genome with a total molecular mass of 150 MDa. The icosahedral capsid is ~927 Å in diameter, not including

the fibers that vary in length from approximately 120 to 315 Å depending on the serotype. The mature adenovirus virion has an icosahedral protein capsid that is composed of three major capsid proteins (240 hexons, 12 penton bases, and 12 fibers) and four additional capsid proteins (proteins IIIa, VI, VIII and IX) (Table 1.1). Its double-stranded DNA genome is packaged inside of the capsid together with four core proteins (V, VII, mu, and terminal protein) and multiple copies of the Adenovirus protease. Cryoelectron microscopy (cryoEM) structures of adenovirus (36, 104), Ad/receptor (22), Ad/antibody (134), and the Adenovirus penton base dodecamer (108) complexes combined with crystal structures for hexon (103), penton base (36), the fiber knob (13, 32, 132, 147), and the fiber shaft (133) have provided a wealth of structural information on Ad, but significant gaps remain in our knowledge of the virus capsid structure. The structure and capsid locations of proteins that play important roles in the virus life cycle (proteins IIIa and VI) and function in capsid integrity (proteins VIII and IX) remain poorly defined.

Table 1.1. Capsid proteins of adenovirus serotype 5.
[1]Size in mature virion.
[2]Trimers of polypeptide II form 240 hexons.
[3]Pentamers of polypeptide III form 12 penton bases.
[4]Trimers of polypeptide IV form 12 fibers, one extending from each penton base.
[5]323 amino acids in Ad35.
[6]35.4 kDa in Ad35.

Name	Number of amino acids	Size (kDa)[1]	Copy Number	Radial Capsid Location	Function
II[2]	951	107.9	720	exterior	Major capsomer.
III[3]	571	63.3	60	exterior	The penton base interacts with the internalization receptor: $\alpha_v\beta_3$ or $\alpha_v\beta_5$ integrins (139).
IIIa	585	63.5	60	exterior/ interior	Viral assembly (9, 25).
IV[4]	581[5]	61.6[6]	36	exterior	The fiber attaches to the host cell via the C-terminal knob.
VI	250	22.1	~ 360	interior	The C-terminus is a cofactor for the adenovirus protease (71, 138). Protein VI regulates the nuclear import of newly translated hexon for viral assembly (141). Protein VI is an endosomal penetration protein (140).
VIII	227	7.6 12.1	~ 120	interior	Capsid stability (65).
IX	140	14.4	240	exterior	Capsid stability. Viral transcriptional activator (98). Protein IX suppresses a cellular antiviral response (98, 99).

Protein IIIa

Protein IIIa participates in viral assembly and maturation as several temperature-sensitive mutants that affect protein IIIa are defective for assembly (9, 25). Protein IIIa is present in 60 copies per virion and is cleaved near the C-terminus by the adenovirus protease during virus maturation to form a mature protein of 570 residues with a molecular mass of 63.5 kDa (Table 1.1) (61, 123). Early immunoprecipitation studies with anti-IIIa antibodies suggested an external location for a least one domain of protein IIIa (34). In contrast, argon plasma etching experiments, cross-linking studies, and proteolytic processing during maturation pointed to an internal position for protein IIIa between the adenovirus capsid and core (34, 82). The first cryoEM difference map at ~25 Å resolution tentatively assigned one domain of protein IIIa to density observed at the facet edge on the exterior of the capsid and a second domain to density on the capsid interior (123).

Fiber

The fiber is a trimer of protein IV, and it extends from each penton base (Table 1.1). The virus attaches to the host cell via the knob at the C-terminus of the fiber. The shaft of the fiber is composed of 15 amino acid consensus repeats that form β-spiral motifs (133). The number of repeats varies among serotypes from 5 and a half in Ad35 to 22 in Ad5. A crystal structure of the Ad12 fiber knob bound to the first Ig-like-domain of CAR shows the fiber interacting with CAR via the side of the knob (5). This tangential arrangement suggests that the fiber must flex to accommodate such an interaction. Many

adenovirus serotypes, including Ad5, contain an insertion in a β-turn in the third repeat from the N-terminus (near the penton base) that could be responsible for the bend observed by negative stain EM of isolated fibers (101). A cryoEM reconstruction of Ad5 shows contiguous fiber density from the capsid until a point corresponding to the third repeat, after which the density becomes diffuse. Conversely, fibers without an insertion in the third repeat like the fiber from Ad37 are more rigid as shown by the fully reconstructed Ad37 fiber (23). The presence of the nonconsensus repeat in the fiber shaft can be used as a predictor of fiber flexibility. The Ad35 fiber also has the nonconsensus repeat, so it is predicted to be flexible.

Protein VI

Protein VI has at least three roles in the adenovirus life cycle. The first biological role of protein VI is that the cleaved C-terminal peptide serves as a cofactor for the adenovirus protease (71, 138). Also, protein VI regulates hexon import into the nucleus during adenovirus assembly (141). Lastly, protein VI participates in endosomal disruption following cell entry (140). Protein VI is present in 369 +/- 34 copies per virion (61) and is cleaved near both the N- and C-termini by the adenovirus protease to form a mature protein of 22.1 kDa (Table 1.1) (61).

Protein VI is thought to be located on the inner capsid surface and has been reported to bind DNA (102). It crosslinks to protein V, a core protein (34) and is degraded with a rate similar to core proteins in Argon plasma etching (82). Direct association between protein VI and hexon has been demonstrated (72). In a 25 Å resolution difference map, protein VI was assigned to internal density associated with the

penton base and peripentonal hexons and was suggested to be a trimer of dimers. A genetic analysis of the Ad5 ts147 temperature-sensitive mutant, which has a defect in nuclear import of hexon, indicates that hexon residue 776 (which is mutated from G to D in ts147) is most likely a key residue in the hexon/protein VI interaction (141). This residue is within the cavity of the hexon trimer and is exposed to the interior of the virion. At non-permissive temperatures the ts147 mutant accumulates newly synthesized hexon trimers in the cytoplasm but the nuclear import of hexon, which is necessary for viral assembly in the nucleus, is impaired (57). In order to efficiently shuttle hexon to the nucleus, it is possible that protein VI associates with hexons regardless of their positions in the assembled virion. Consideration of the capsid protein stoichiometry in the mature virion indicates 1 to 2 copies of protein VI per hexon trimer, with ~369 copies of protein VI to 240 hexon trimers per virion. A high resolution structure of adenovirus may be necessary to determine whether protein VI associates preferentially with peripentonal hexons as proposed in the first adenovirus difference map or binds equivalent sites in other hexons in the asymmetric unit.

Protein VIII

The location of protein VIII within the virion is better characterized than proteins IIIa and VI. Protein VIII is thought to stabilize the capsid since a temperature-sensitive mutant with a substitution mutation in protein VIII is more thermolabile than wildtype virus (65). A cryoEM analysis of adenovirus particles at ~10 Å resolution has provided a convincing assignment for protein VIII on the inner capsid surface (36). Protein VIII is known to be associated with hexons and is present in ~120 copies per virion (Table 1.1)

(123). Comparison of cryoEM structures of wild type Ad5, and a mutant virus lacking protein IX that had undergone single or multiple freeze/thaw cycles, enabled Fabry *et al*. to observe that there are two similar density regions per asymmetric unit on the inner capsid surface (36). These density regions were approximately the correct volume for protein VIII and in the correct copy number, which helped to support the assignment. Mass spectrometry indicates at least two adenovirus protease cleavage sites in the protein VIII precursor, leading to 7.6 and 12.1 kDa protein fragments in the mature virion (61). The precursor form of protein VIII has been found to be useful in reverse-phased HPLC analyses to quantitate the percentage of empty capsids in adenovirus preparations (125, 135). This implies that mature adenovirus virion contains only the two cleaved fragments of protein VIII, with no remaining copies of the precursor form of protein VIII.

Protein IX

Protein IX is known as the "cement" that holds the capsid together. The main locations of protein IX within the virion were deduced by scanning transmission electron microscopy (STEM) of capsid dissociation fragments (40). Protein IX is present in 240 copies per virion (Table 1.1), and four trimers of protein IX are known to stabilize the group of nine hexons in the center of each icosahedral facet. Protein IX is a relatively small capsid protein of 14.4 kDa, but recently it was shown to have at least two additional roles in adenovirus biology: protein IX is a transcriptional activator that enhances viral replication (98) and late in infection protein IX sequesters promyelocytic leukemia (PML) protein in clear amorphous nuclear inclusions (98, 99). PML proteins are the organizers of nuclear domains or PML bodies that participate in cellular antiviral responses (19).

Also, protein IX has gained prominence as a platform for ligand addition for the purposes of vector retargeting and fluorescence labeling (90). Thus, it is of interest to further investigate the location of protein IX and the other components within the adenovirus capsid.

Aim I

In Aim I, a subnanometer resolution cryoEM structure of adenovirus is determined. Ad35F was selected for this structural study due to its promise as a gene therapy vector targeted to hematopoietic cells and also because its short fiber is advantageous for cryoEM imaging. After using the hexon and penton base crystal structures to computationally remove hexon and penton base density from the EM map, the structure of the fiber and proteins IIIa, VI, VIII, and IX can be better characterized. In favorable cases, ~10 Å resolution is sufficient to resolve bundled α-helices. In the 9 Å resolution structure of Ad35F presented in Chapter III we resolved one α-helix on the capsid surface that was assigned to the C-terminal domain of protein IX. On the basis of this structural study we postulated that protein IX might have two alternate conformations either binding on the capsid surface or extending away from the capsid. This model is consistent with the tolerance of the C-terminus for inserted ligands and its potential use in vector retargeting. Well-defined density was assigned to the conserved N-terminal domain of protein IX required for incorporation into the virion.

Adenovirus cell entry and life cycle

In modifying vector tropism, a detailed understanding of the cell entry process is required so that attachment to the host cell, internalization, and transfer of the genome to the nucleus remain intact. Adenovirus particles attach to host cells via characteristic fibers noncovalently extending from the penton base at the vertices. $\alpha_v\beta_3$ or $\alpha_v\beta_5$ integrins cluster at the penton base and initiate endocytosis (139). The fiber is thought to be released while the virion is still at the cell surface (81). Further viral dissociation occurs in the acidified environment of the endosome (47) and partial disassembly of the adenovirus capsid is required for membrane disruption (140). Protein IIIa is likely to play a role in cell entry as it is one of the earliest components released from the virion and one of the most thermally labile capsid components (47, 96, 140). Upon exposure to acidic pH, approximately 25% of hexons, the majority of penton bases, protein IIIa, and 80% of the internal capsid protein VI dissociate from the virion. Protein VI possesses membrane lytic activity (140). The conformational changes that the adenovirus capsid must undergo in order to release protein VI are not well understood. Following membrane penetration, the partially disassembled virion travels to the nucleus via microtubules (62) and docks at the nuclear pore complex (NPC) where further disassembly occurs and the genome is transported to the nucleus (45).

Transcription begins once the viral genome is imported into the nucleus. Early gene expression prepares the cell for viral replication. The immediate-early expressed products of the E1A and E1B genes activate transcription, induce S phase, and block apoptosis. The expression of the remaining early genes (E2, E3, E4, and VA) shuts down host mRNA transport and activates viral DNA replication. In the cytoplasm, the E3 gene

products result in immune evasion. Late transcripts produce the structural proteins needed for progeny virions (116). Viral assembly occurs in the nucleus, and protein VI regulates the nuclear import of hexons for assembly into virions (141). Viral DNA is packaged into preformed capsids with the assistance of protein IIIa and scaffolding proteins; however, the scaffolding proteins are released by the time mature virions are formed (44). The adenovirus cysteine protease cleaves numerous capsid and core proteins, thereby producing mature viral proteins. The proteolytic processing of precursor capsid and core proteins is necessary to produce infectious particles. Late in infection, high levels of the adenovirus death protein induces cell lysis and release of progeny virions (127). The lack of an atomic resolution structure of an intact adenovirus virion has limited our understanding of the adenovirus cell entry process and life cycle and therefore hindered further vector development.

Aim II

The process of acid-induced disassembly and its relationship to protein VI release and lysis of the endosome may be further probed with a higher resolution structure of adenovirus. As the resolution of a density map improves α-helices resemble rods of density. Although ~10 Å resolution might be sufficient to resolve α-helices in favorable cases, many α-helical density rods are not fully resolved until ~7 Å resolution (56). Thus, in Aim II, a higher resolution structure of Ad35F is determined with the goal of identifying α-helices in the capsid proteins. Visualization of α-helices assisted the interpretation of capsid density remaining after the hexon and penton base density was

segmented and helped address how the locations of capsid proteins are linked to low pH-induced capsid disassembly during cell entry.

Selection of the technique for structure determination

Structure determination by X-ray crystallography or solution state NMR works well for proteins that crystallize or for proteins that are less than 100 kDa in molecular weight, respectfully. Crystal structures have been solved for hexon (103), penton base (36), and portions of the fiber (13, 32, 133, 147); however, at least two of the four capsid components without atomic resolution structures are not soluble outside of the virion, which makes structure determination of individually expressed or purified capsid proteins difficult. Solving a high resolution structure of the intact adenovirus virion may be the only way to structurally characterize proteins IIIa, VI, VIII, and IX. Thus far, crystallography has not provided an atomic resolution structure for a virus as large as adenovirus, which has an icosahedral capsid 930 Å in diameter and a molecular mass of 150 MDa. The largest viral structure solved by X-ray crystallography to date is simian virus 40 with a diameter of 500 Å and molecular mass of 14 MDa, which is approximately one tenth the mass of adenovirus (119).

CryoEM is a powerful technique for structure determination of macromolecular complexes like adenovirus that are not amenable to other types of structural analyses. For cryoEM imaging, a small amount of the purified biological complex is flash frozen on a holey carbon grid in a thin layer of ice (2). The temperature is dropped below –140 °C so rapidly that vitreous or glass-like ice forms. Vitreous ice is not ordered like hexagonal ice, so the structure of the specimen is preserved in a near native-like state. There is no need

to crystallize the specimen because the random orientations of particles when frozen are advantageous for image processing.

Major technological advances in imaging and computation have made subnanometer (<10 Å) resolution cryoEM structures attainable. These advances include the commercial availability of the FEI Polara electron microscope, which is designed specifically for cryo work and has a specimen stage that is significantly more stable than electron microscopes with third party cryo-holders. With the Polara microscope specimen drift, which typically causes loss of high resolution information in cryo-micrographs, is kept to a minimum. The Polara microscope also has a highly coherent electron beam produced by a field emission gun (FEG), and a high accelerating voltage (300kV), both of which contribute to producing cryo-electron micrographs with strong high resolution information. Advances in image processing include the ability to process large 3D volumes (up to 750^3 voxels) and improved schemes for correcting the effects of the contrast transfer function (CTF) of the electron microscope.

CHAPTER II

MATERIALS AND METHODS

Gene delivery assay

The Ad35F vector, with the Ad35 fiber, and the control Ad5F vector, with the Ad5 fiber, were prepared as described previously by Smith *et al* (117). The human acute monocytic leukemia cell line THP-1 was grown in RPMI 1640 medium supplemented with 10mM HEPES pH 7.55, 0.1mM non-essential amino acids, 4mM L-glutamine, 100 units/ml penicillin G sodium, 100 μg/ml streptomycin sulfate and 10% fetal calf serum. Cells were plated at $4x10^4$ cells/well in a 96 well plate. Ten microliters of virus were added to the 100 μl culture volumes in triplicate. After an overnight incubation at 37ºC, β-galactosidase activity was assayed using the Tropix Gal-Screen chemiluminescent reporter gene assay system (Applied Biosystems, Bedford, MA) according to manufacturer's instructions using an equal volume of Buffer A for mammalian cells. Luminescence was read using an EG & G Berthold Microplate Luminometer LB96V (Perkin Elmer-Wallac Inc., Gaithersburg, MD) with a 1 sec time interval.

Vitrification

To remove the stabilizing agent, CsCl, small aliquots of purified Ad35F were dialyzed in 0.05 M Tris buffered saline (TBS) at pH 8.0 for 1 hour three times. pH 8.1 was used to avoid pH-induced capsid dissociation. CryoEM grids were prepared using standard techniques (2). Briefly, 4 μl of purified, dialyzed Ad35F was applied to R2/4

Quantifoil holey carbon grids (Quantifoil Micro Tools GmbH) and flash frozen in ethane slush using a Vitrobot (FEI, Eindhoven, The Netherlands). Prior to plunging, a carbon layer was added to each side of the Quantifoil grid with the plastic film dissolved in between carbon layer applications to strengthen the grid support layer for imaging with the TF30 Polara electron microscope (FEI Company, The Netherlands). To reduce beam-induced movements of the grid, the support layer was conditioned under an intense beam in a CM12 electron microscope overnight after the first carbon layer application.

CryoEM imaging

Data collection was performed on an FEI Polara (300kV, FEG) electron microscope at liquid nitrogen temperature and 300kV with an UltraScan 4k x 4k CCD camera (Gatan Inc., Pleasanton, CA). Datasets 1 and 3 were collected in the nanoprobe imaging mode while dataset 2 was collected under the microprobe mode. Datasets 1 and 2 were collected with an absolute magnification of 253,654X, and dataset 3 with 397,878X. Leginon (124) was used for semi-automatic collection of dataset 2. Datasets 1, 2, and 3 contained 1511, 4657, and 2015 micrographs, respectively. The microscope magnifications were determined to within +/-0.5% with a negatively stained TMV grid and later refined to within +/-0.25% during image processing for maximum correlation between the datasets and between the hexon coordinates (103) and the cryoEM density. The refined Angstrom-to-pixel ratio is 0.589 Å for datasets 1 and 2 and 0.375 Å for dataset 3 on the molecular scale without binning.

Three-dimensional image processing

The entire dataset consisted of 6,880 particle images. Dataset 1 particle images (2214) were selected manually with QVIEW (112) and an automated particle picking program, Virus (1), was used for datasets 2 (2724 images) and 3 (1919 images). The automated routine CTFFIND3 (76) was used to determine the initial defocus and astigmatism parameters for each particle image.

Initial refinement of Ad35F data for Aim I

IMAGIC-5 reconstruction software was used for generation of the Ad35F structure at 8.9 Å resolution. This software does not include a mechanism to correct astigmatic particle images. Thus, particle images from dataset 1 with the greatest astigmatism (with a focal difference greater than 1,500 Å along the axis of astigmatism) were rejected from further processing. Through this mechanism the best 63% of the data was selected. The EMAN routine, applyctf, was used to flip phases using the defocus values determined by ctffind3, and thus partially correct for the contrast transfer function (CTF) of the microscope (66). An inverse temperature factor was not applied to the data.

The IMAGIC-5 image processing software package was used to iteratively refine the centers of the particle images, determine angular orientations, and calculate three-dimensional reconstructions (131). For any cryoEM dataset the best theoretically attainable resolution is twice the pixel size, termed the Nyquist limit. In practice, resolution better than three times the pixel size is rarely achieved. To achieve subnanometer resolution, or ~10 Å resolution, a pixel at least as fine as 3 Å should be processed. To speed up computation, the initial angular orientations were determined for

particle images binned to 250^2 pixels (4.7 Å/pixel) and subsequently refined using particle images binned to 500^2 pixels (2.4 Å/pixel). In the final round of image processing, particle images binned to $1{,}000^2$ pixels (1.2 Å/pixel) were used for refinement and then binned to 500^2 pixels for calculation of the three-dimensional reconstruction (with 500^3 voxels, 2.4 Å/voxel).

Phase residual-based sorting, similar to that performed by Borgnia *et al*. (8), was employed to select the best subset of particle images. Several times during refinement a phase residual was calculated between each particle image and its corresponding reprojection, which was calculated by projecting the reconstruction along the orientational angles determined for that particle image. The differential phase residual function in IMAGIC-5 was used for this purpose (130, 131). The data set exhibited a bimodal distribution of phase residuals. The subset of particle images with lower average phase residual values (82˚ to 94˚) were selected and independently reconstructed. This selection process improved the resolution by nearly 0.5 Å in the final round of refinement. The final reconstruction is based on 964 particle images. The resolution of the final reconstruction was assessed to be 8.9 Å by the Fourier shell correlation (FSC) method with the 0.5 correlation threshold (88). Inner and outer soft Gaussian masks, at radii 277 and 465 Å respectively, were applied to the reconstruction so that the resolution determined would correspond to just the icosahedral virus capsid (121). This masking procedure removed the distal end of the fiber; however it did not remove any of the ordered density on the inner capsid surface.

High resolution refinement of Ad35F data for Aim II

Similar to the processing of images in Aim I, particle images were binned to 250^2 pixels for the initial rounds of refinement, and subsequently binned to 500^2 pixels for additional processing with a ~2 Å pixel size (2.36 Å for datasets 1 and 2, and 2.25 Å for dataset 3), and later binned to 750^2 pixels for image processing with a 1.5 Å pixel size. The 9 Å resolution reconstruction of Ad35F (104) was used as a starting model for image processing with the FREALIGN v.7 program (49, 143). The FREALIGN code was modified to allow the input of angular and translational step sizes for refinement and the input of externally determined particle centers to aid in the initial determination of the orientational angles. The IMAGIC-5 (131) centering algorithm, Align, was used to determine initial particle centers.

Multiple rounds of refinement were performed with FREALIGN, first with refinement of the five orientational parameters (3 angular and 2 translation), and later with additional refinement of the defocus and astigmatism parameters. The orientational angles were refined to within 0.0005° and X- and Y-centers to within 0.0025 pixels. No inverse temperature factor was applied to the data. Frealign can account for particle astigmatism in the reconstruction routine, so all the data was included for the purposes of orientational refinement. A phase residual cutoff of 85.5° was utilized to include only images with the highest correlation to the 3D reconstruction. During refinement the particle images were masked with an outer radius of 463 Å, which effectively cut off the fibers. For FSC resolution assessment, inner and outer masks at radii 300 Å and 463 Å were applied to prevent the inclusion of density from the disordered core and flexible fibers. Refinement was performed on an SGI Altix computer with 16-processors and

32GB of RAM, and on the Vampire 64-bit OS linux cluster at Vanderbilt. The final reconstruction was calculated from 2,630 particle images.

Difference map generation

Difference mapping analyses were performed on the 6 and 9 Å resolution structures of Ad35F, but for Aim II a C-shell script (Gen_diffmap) was written to complete the difference mapping analysis in an automated fashion. First Gen_diffmap directed the docking of the hexon (103) and penton base (157) coordinates into subvolumes of the cryoEM map using the automated fitting tool, CoLoRes, in the Situs package (15). Gen_diffmap selected the best solutions corresponding to the four unique hexon positions within the asymmetric unit. The coordinates for a penton base monomer were extracted from the best solution for the penton base. Icosahedral symmetry was applied to the asymmetric unit to yield a pseudoatomic capsid of 240 trimeric hexons and 12 pentameric penton bases. The pseudoatomic capsid was converted into a density map and filtered to 6.9 Å resolution (8.9 for Aim I) with the pdb2mrc routine in EMAN v1.7(66), normalized to the cryoEM density with the IMAGIC command Norm-Variance (131), and subtracted from the cryoEM map with the IMAGIC command Subtract-Images in order to generate a difference map. The difference map contained density for the minor capsid proteins IIIa, VI, VIII, and IX, as well as small part of fiber associated with the penton base and regions of hexon and penton base missing from the crystal structures. Lastly the Gen_diffmap script extracted one eighth of the cryoEM reconstruction, the pseudoatomic map, and the difference map for visual inspection with

the UCSF Chimera package (92). Molecular graphics images were produced using Chimera and AVS5 (Advanced Visual Systems, Inc.)

Volumetric analysis of capsid proteins

When the 9 Å cryoEM reconstruction is contoured at 1σ, the adenovirus capsid surface appears continuous and there is essentially no noise. The volumes of various extracted density regions were assessed with a Fortran program (density_mult), which sums the volume of all of the voxels in the density region that have a density value equal to or greater than the specified contour level. For volume analysis of three regions in the structure determined in Aim I: the internal density associated with penton base; the large density lobe at the facet edge on the outer capsid surface; and the N-terminal third of protein IX, a contour level of 1 σ was used.

Secondary structure prediction

The PsiPred server(12) was used to predict secondary structure for the sequences of the full-length precursor forms of the Ad5 proteins IIIa, VI, VIII, and IX, as well as the N-terminal 51 residues of the Ad5 penton base.

Density segmentation, fitting of polypeptide models, and identification of protein interaction residues

The Volume Eraser function of Chimera (92) was used for manual density segmentation. Canonical α-helices were built using Swiss PDB Viewer (50) for each length of helix 10 residues or longer predicted by PsiPred for protein IIIa. The α-helices

and 7-residue polypeptide models were manually positioned with Chimera into density rods in the cryoEM density.

Potential protein-protein interaction residues between penton base and hexon in the pseudoatomic map and proteins IIIa in the cryoEM difference map were evaluated with Chimera (92). Protein IIIa density was displayed together with either the penton base or hexon coordinates. Penton base or hexon residues within 10 Å of protein IIIa were selected and saved to an output file. Some residues were eliminated as potential interaction residues because other residues blocked their 'line-of-sight' to protein IIIa.

CHAPTER III

CRYO-EM STRUCTURE AT 9 Å RESOLUTION OF AN ADENOVIRUS VECTOR TARGETED TO HEMATOPOIETIC CELLS

Aim

Apart from the density assignments in the first adenovirus difference map at 25 Å resolution (123), very little is known structurally about capsid proteins IIIa, VI, VIII, and IX. The goal of Aim I was to characterize the capsid components without known atomic resolution structures by determining an intermediate resolution cryoEM structure of adenovirus. A subnanometer resolution structure should allow a detailed volumetric analysis of capsid density remaining after the hexon and penton base are computationally removed. This chapter was published as a manuscript in the Journal of Molecular Biology (104).

Experimental design

The vector (Ad35F) used in the study is composed of the Ad5 capsid pseudotyped with the Ad35 fiber. Adenovirus fibers are composed of a variable number of β-spiral motifs called β repeats. The Ad35F vector is ideal for a cryoEM study because the Ad35 fiber has only 5 and a half β-spiral repeats compared to 22 β-spiral repeats in Ad5. This shorter fiber should allow a thinner layer of ice around the adenovirus particle, which should in turn produce cryo micrographs with the best signal-to-noise ratios. Ad35F is interesting for structural analyses since it shows promise in vector retargeting to non-CAR expressing cells including hematopoietic cells. Non-hexon and non-penton base

density is visualized by a difference mapping approach where a pseudoatomic capsid of 240 hexons and 12 penton bases is subtracted from the cryoEM map. By volumetric analysis together with modeling studies, the minor capsid components are further characterized.

Adenovirus vector gene delivery to hematopoietic cells

Most adenoviral vectors used in gene therapy applications are derived from the common Ad5 serotype, which infects a broad range of cells and normally causes a mild respiratory infection. A panel of adenovirus vectors based on the Ad5 serotype was generated with various mutations in the fiber and penton base and these were screened for *in vivo* gene transfer in mice (117). One vector, Ad35F, was selected for a high resolution cryoEM structural analysis on the basis of its short fiber and gene delivery potential. This vector is composed of the entire icosahedral capsid of Ad5 with the exception of the Ad35 fiber. Ad35F increased gene delivery to human monocytic cells ~100 fold compared to the same vector with the Ad5 fiber (Ad5F) (Figure 3.1). Ad5 normally utilizes CAR for host cell infection, while Ad35 uses CD46. This result supports the emerging concept that the CD46-utilizing fibers play a major role in promoting infection of human hematopoietic cells (114).

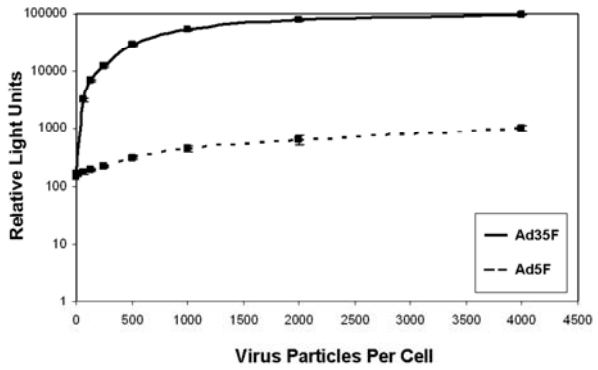


Figure 3.1. Gene delivery by Ad35F to a leukemia cell line. THP-1 cells were incubated overnight in triplicate wells ($4x10^4$ cells/well) of a 96 well plate without or with increasing concentrations of Ad35F (solid) or Ad5F (dashed) containing the β-galactosidase gene. Infection, determined by increased β-galactosidase activity, was assayed using the Tropix Gal-Screen chemiluminescent reporter gene assay system and read using an EG & G Berthold Microplate Luminometer LB96V. Data points are means of triplicate wells ± SD. Figure reprinted from (104) with permission from Elsevier.

CryoEM structural analysis of Ad35F

Relatively few virus structures have been determined to better than 10 Å resolution by cryoEM methods and adenovirus is one of the largest of these (155). In order to achieve high resolution for adenovirus, cryoelectron micrographs were collected on an FEI Polara microscope and image processing was performed on a parallel computing cluster. The final data set of the best 964 Ad35F particle images produced a reconstruction at 8.9 Å resolution (Figure 3.2). The diameter of the icosahedral capsid measured along a 5-fold symmetry axis is ~927 Å. If the protruding fiber is considered including the weak density from the fiber knob, then the diameter along a 5-fold axis is ~1,191 Å. Evaluation of the cryoEM density for this large and complex virus was greatly aided by several factors: the crystal structures of the Ad5 hexon (103) and Ad2 penton base (157) with 98% sequence homology to the Ad5 penton base, an earlier cryoEM difference map analysis of Ad2 (123), and knowledge of the stoichiometry of the capsid proteins (61).

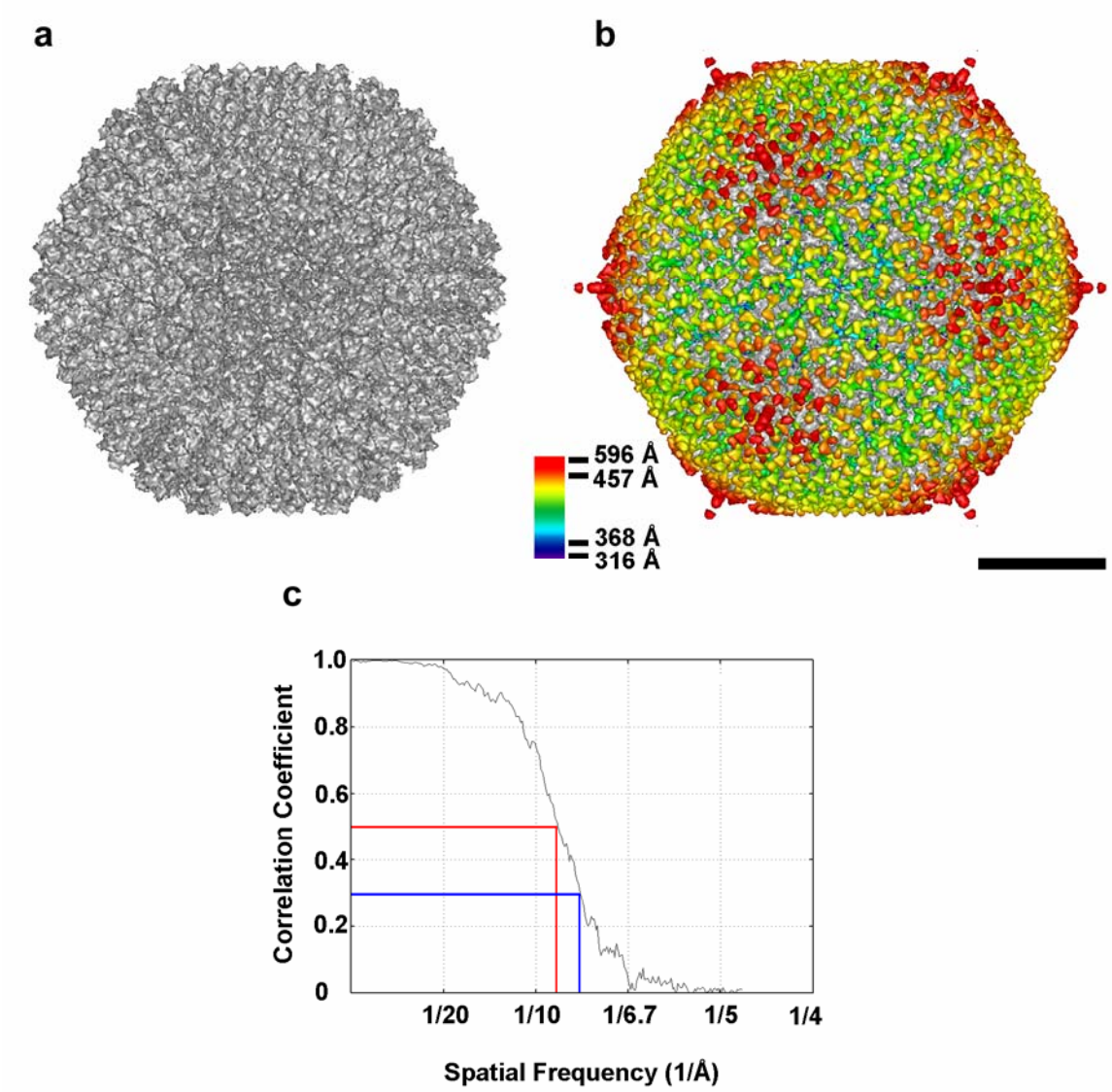


Figure 3.2. CryoEM structure of Ad35F. **(a)** The pseudo-atomic capsid formed by 240 copies of the trimeric hexon and 12 copies of the pentameric penton base shown filtered to 9 Å resolution and viewed along an icosahedral 3-fold axis. **(b)** CryoEM difference density (radially color coded) together with the pseudo-atomic capsid (gray) in the same view as in part (a). Bar, 250 Å. **(c)** A Fourier shell correlation (FSC) plot indicating a resolution of 8.9 Å at the 0.5 threshold (red) and 8.1 Å at the 0.3 threshold (blue). Figure reprinted from (104) with permission from Elsevier.

CryoEM density for hexon and associated internal density

The Ad5 hexon crystal structure is missing 60 residues due to disorder, and all but 9 of these are on the exposed top of the hexon (103). Figures 3.3a and 3.3b show cryoEM density for one of the three towers at the top of a hexon next to an icosahedral 3-fold axis. The crystallographic coordinates for this region are also displayed with the missing residues and sidechain atoms in this region indicated. In the difference map calculated between the Ad35F cryoEM reconstruction and the pseudo-atomic capsid, there is difference density at the top of each hexon. In addition to the difference density at the top of the hexon, there are also small patches of difference density on the hexon sides that match regions where only sidechain atoms are missing from the crystal structure.

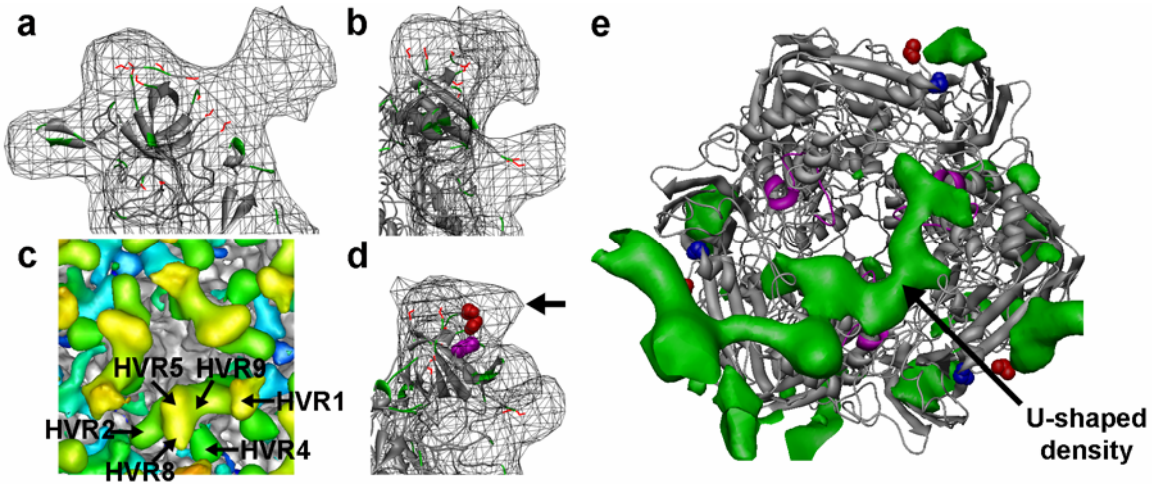


Figure 3.3. CryoEM hexon density. **(a)** Side view of one tower from a hexon next to an icosahedral 3-fold axis. The cryoEM density (mesh) encloses the tower coordinates from the docked Ad5 hexon crystal structure (PDB 1P30) (103) (ribbon). The residues in the tower region that are missing from the crystal structure are represented by red gaps, and the residues with missing sidechain atoms are represented in green. **(b)** A 90˚ rotated view of the cryoEM density and docked hexon coordinates colored as in part (a). **(c)** CryoEM difference density at the top of the same hexon shown radially color coded with labels indicating the general locations of hypervariable regions (HVRs) 1, 2, 4, 5, 8, and 9 that contribute to the difference density. **(d)** The cryoEM hexon density (mesh) and the Ad5 hexon tower coordinates (ribbon) colored as in parts (a) and (b) with the residues near the epitope of a neutralizing antihexon antibody (9C12). Epitope proximal residues 178 and 179 from the DE1 loop (magenta) and residues 430 and 437 from the FG1 loop at either end of a gap (red) are shown as spheres. The hexon is rotated slightly more that 90˚ with respect to part (a) so that the density protrusion (arrow) near the 9C12 epitope is clearly visible. **(e)** Bottom view of the hexon trimer coordinates docked into position H2 displayed with the associated internal cryoEM difference density (green). The U-shaped difference density lining one side of the hexon cavity is indicated. The first and the last ordered hexon residues near the N- and C-termini are shown as blue and red spheres, respectively. The residues implicated in protein VI binding (aa 770 - 787) are shown in magenta. Figure reprinted from (104) with permission from Elsevier.

Many of the gap regions in the hexon crystal structure correspond to hypervariable regions (HVRs) between hexons of various adenovirus serotypes, which include type-specific epitopes (28, 103). We have labeled the difference density at the top of the hexon in terms of these HVRs (Figure 3.3c). The missing region (aa 136-164) that roughly corresponds to HVR1 (aa 138-166) is quite long. Therefore HVR1 might cover a large surface area on the top of the hexon. At this resolution it is not possible to precisely delineate the boundaries of the HVRs. However, the wide wing-shaped hexon density observed in the cryoEM difference density clearly includes missing residues and sidechain atoms from HVRs 1, 2, 5, 8 and 9. In addition a separate lobe of density is observed on side of the hexon tower corresponding to missing residues (aa 251-256), which overlap with HVR4 (aa 253-257).

The hexon difference density most likely includes the epitope recognized by the 9C12 neutralizing monoclonal antibody, which has been visualized in complex with Ad5 by cryoEM (134). In this previous study, hexon residues 178 and 179 as well as residues from the FG1 loop were identified as close to the center of the epitope. However the moderate resolution (21 Å) of the cryoEM structure of the Ad5/9C12 complex did not allow a precise identification of the epitope. The 9 Å resolution Ad35F cryoEM structure shows a density protrusion near the epitope proximal residues that is likely to contain the actual epitope for 9C12 (Figure 3.3d). This density protrusion is most likely formed by residues from the gap of aa 431-436, which includes HVR8 from the FG1 loop, and possibly by residues from the long gap of aa 136-164, which includes most of HVR1.

On the inner capsid surface a significant amount of difference density was observed below each hexon (Figure 3.3e). The hexon crystal structure is missing 4

residues at the N-terminus and 5 residues at the C-terminus, both of which are located near the bottom of the hexon, however the difference density is much larger than expected for these residues. A U-shaped region of difference density appears to line one side of the hexon cavity and connects with a large density lobe below the interface between several hexons. The U-shaped density comes within 10 Å of the hexon residues thought to be involved in interactions with protein VI (72, 141).

Penton base integrin-binding loop, variable loop, and associated internal density

The Ad2 penton base crystal structure is missing ~80 residues in the integrin-binding RGD loop due to disorder (157). Since the Ad2 and Ad5 penton base sequences are highly homologous (98% identity), the Ad2 penton base crystal structure fit well within the cryoEM reconstruction of Ad35F, which has the Ad5 penton base. CryoEM density from five RGD loops is clearly visible protruding above the penton base (Figure 3.4a). In addition, five copies of a second penton base surface loop are also observed in the difference map. This region corresponds to the variable loop on the side of the penton base (aa 142-169 in Ad2 and Ad5) that is present in the crystal structure as an extended β ribbon. The observation of cryoEM density for this surface loop, despite the fact that it is in the crystal structure, indicates that this loop is flexible.

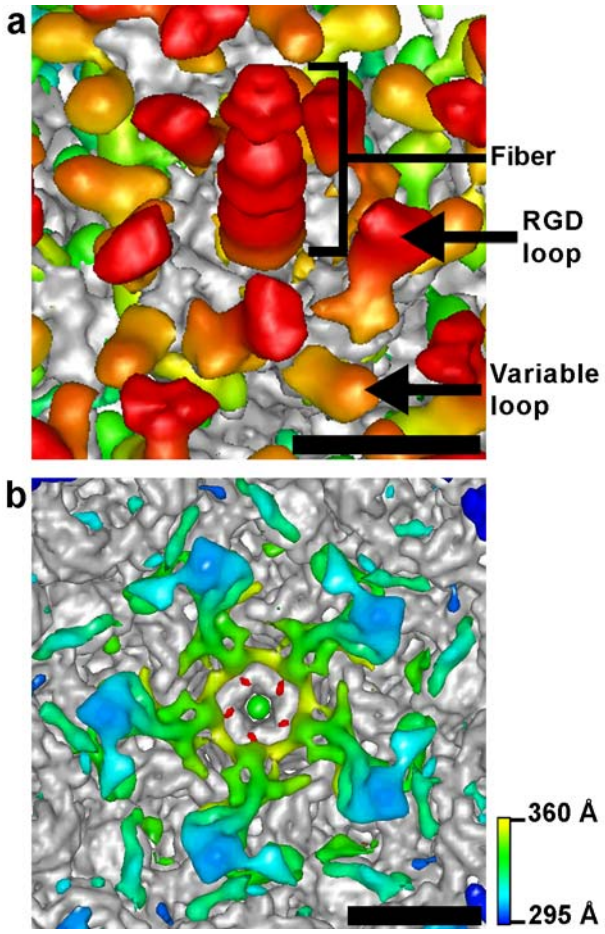


Figure 3.4. CryoEM penton base density. **(a)** Tilted view of the penton base and fiber difference density. Density regions corresponding to the fiber, penton base RGD loop, and variable loop are indicated. The difference density is radially color coded as in Figure 3.2a. **(b)** Internal capsid density below the penton base. The N-terminal five residues (aa 52-56) in the penton base crystal structure (PDB 1X9T) (157) are shown as red density, which is only partially visible at the bottom of the penton base. Bars, 50 Å. Figure reprinted from (104) with permission from Elsevier.

In addition to the density on the outer surface of the penton base, there is also cryoEM difference density below the penton base on the inner capsid surface (Figure 3.4b). Five density spokes extend ~50 Å into the viral core. The penton base crystal structure was determined for an N-terminal truncation missing the first 48 residues. The first ordered residue present in the crystal structure is at the bottom of the penton base. Thus, the cryoEM density observed below the penton base probably includes the missing N-terminal tails. Volume analysis indicates that an additional ~20 kDa of mass is associated with each penton base N-terminal tail.

Ad35 fiber flexibility

CryoEM difference density is observed for the fiber at each capsid vertex. Negative stain EM and fiber sequence alignment suggested that the fibers of some adenovirus serotypes are flexible due to variations in the fiber shaft repeat sequence (24). This concept was supported by the crystal structure of a portion of the Ad2 fiber shaft, which further elucidated the residues important for the fiber shaft triple β-spiral fold (133). Analysis of the Ad35 fiber sequence revealed five and a half repeats; however, repeat number 3 deviates from the consensus and is predicted to be flexible (Figure 3.5a). The observed cryoEM density for the Ad35 fiber is in agreement with this prediction as the fiber shaft density appears discontinuous and the fiber knob density is weak (Figure 3.5b and 3.5c).

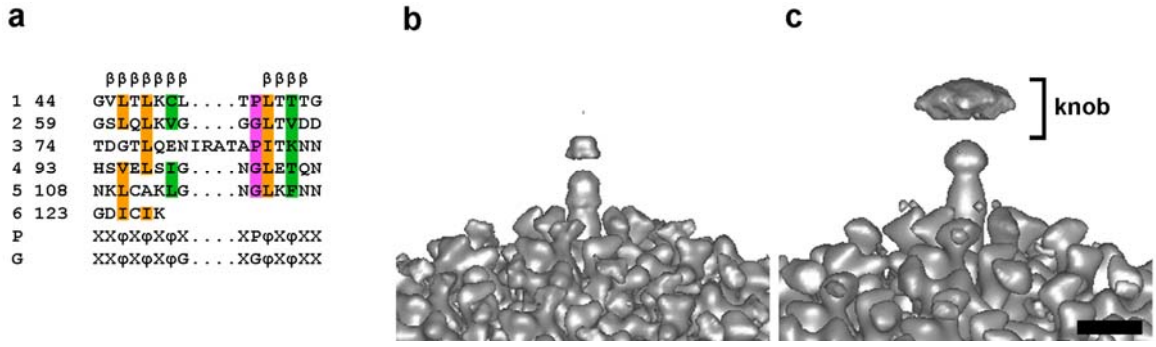


Figure 3.5. Ad35 fiber density. **(a)** Alignment of the five and a half repeats in the Ad35 fiber shaft colored as by van Raaij *et al*.(133): with the conserved hydrophobic residues in the core in orange; the residues forming an external patch in green; and the conserved proline or glycine in purple. The two consensus patterns are shown below. **(b)** Side view of the Ad35F vertex region including a penton base and fiber. The fiber shaft appears discontinuous presumably because repeat 3 is flexible. **(c)** The Ad35F vertex region shown filtered to 24 Å and with a lower isosurface to include weak density for the fiber knob (bracket). Bar, 50 Å. Figure reprinted from (104) with permission from Elsevier.

The crystal structure of the Ad2 penton base/fiber peptide complex shows that 11 residues (aa 10-20) of the N-terminal fiber tail lie in a groove on the top of the penton base (157). Therefore, the homologous region in the Ad35 fiber would likely contribute no appreciable length to the reconstructed fiber. The length of a triple β-spiral repeat is 13 Å (133). There is currently no atomic resolution structural information for the sequence between the penton base binding peptide and the first repeat (aa 21-43 in Ad35). Since this region is approximately the same number of residues as in one fiber shaft repeat, we estimate that it has approximately the same height, 13 Å. For the half repeat at the C-terminal end of the shaft we estimate 6 Å. Thus, we predict that the reconstructed Ad35 fiber shaft should be 84 Å long, in excellent agreement with the measured length of 85 Å.

When the reconstruction is filtered to 24 Å and shown with a lower isosurface to include weak density, density is observed for the fiber knob (Figure 3.5c). However, the knob appears wider than expected from the knob crystal structures (13, 32, 132, 147). This widening is likely due to the flexible Ad35 fiber shaft that results in smeared, weak reconstructed density for the knob. The height of the fiber knob, as measured from the knob crystal structures, is ~50 Å. Therefore we would predict a length of 134 Å (84 Å for the shaft plus 50 Å for the knob) for the full length Ad35 fiber. This is in good agreement with the fiber length of 132 Å as measured in the 24 Å filtered Ad35F reconstruction.

The adenovirus capsid has a well characterized symmetry mismatch between the trimeric fiber and the pentameric penton base. In the icosahedral cryoEM reconstruction, 5-fold symmetry has been imposed on the trimeric fiber effectively blurring the fine features of the fiber. So although the crystal structure of the fiber shaft revealed a

distance of 13 Å and a rotation of ~50° from one repeat to the next within a single fiber polypeptide (133), we would not expect to observe these features in the reconstruction.

Additional hexon associated density

In an earlier cryoEM difference map analysis of Ad2 at ~25 Å resolution, large discrete regions on the outer capsid surface were assigned to protein IIIa, with two copies along each icosahedral 2-fold edge for a total of 60 copies within the virion (123). These same density lobes are apparent in the 9 Å resolution difference map between hexons of two different facets (Figure 3.6a). The HVR4 loop of one neighboring hexon interacts with this large density lobe (Figure 3.7a). Volume analysis indicates that the large density lobe corresponds to ~27 kDa, which is approximately half of the molecular mass of the mature form of protein IIIa (63 kDa) (61).

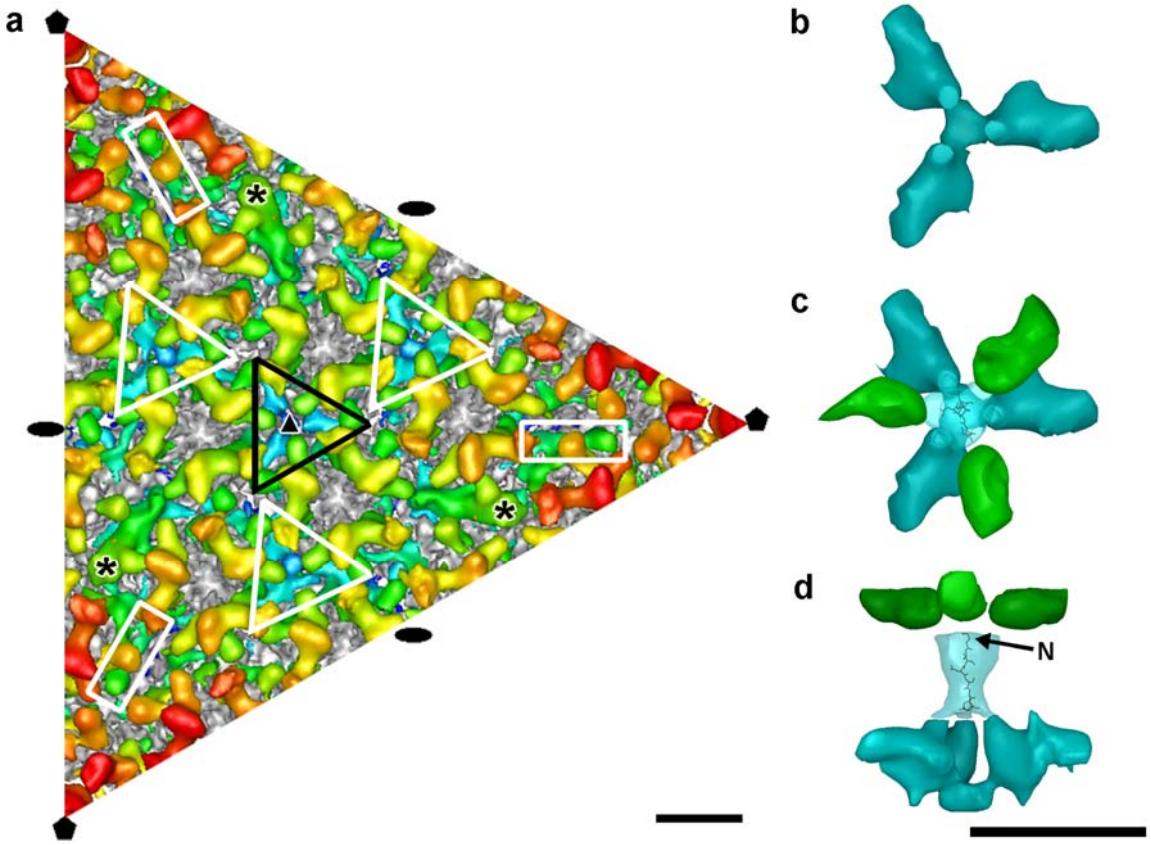


Figure 3.6. N-terminal domain of protein IX. **(a)** A region of the cryoEM difference density that roughly corresponds to an icosahedral facet. A small portion of fiber (red) is visible in the three corners of the image. Trimeric density assigned to the N-terminal domain of protein IX (cyan) is outlined at the icosahedral 3-fold axis (in black) and at three local 3-fold axes (in white). White rectangles outline the density assigned to the C-terminal region of protein IX (green), which is partially occluded by the peripentonal hexons. A large density lobe near the facet edge, which is tentatively assigned as a significant portion of polypeptide IIIa, is marked with an asterisk. Crystallographic symbols indicate the icosahedral 2-fold (black oval), 3-fold (black triangle), and 5-fold (black pentagon) symmetry axes. **(b)** Top view of the trimeric density assigned to three copies of N-terminal domain of protein IX extracted from the icosahedral 3-fold region. **(c)** Top view of the same density together with three nearby hexon HVR2 regions (green) and weak density (transparent cyan) extending outward from the trimeric density to the hexon HVR2 regions. **(d)** Side view of the protein IX and hexon density shown in part (c). The weak density is fitted with a model for the N-terminal 7 residues of protein IX. Bars, 50 Å. Figure reprinted from (104) with permission from Elsevier.

CryoEM density for protein IX

Mass spectroscopy has confirmed that there are ~240 copies of protein IX within the adenovirus capsid (61). Density was assigned to protein IX by cryoEM difference mapping (123) on the basis of an earlier Scanning Transmission Electron Microscopy (STEM) study (40). Trimeric protein IX density was observed in the 9 Å resolution cryoEM difference map in the same four locations per facet as previously observed (Figure 3.6a). The trimeric density region at the icosahedral 3-fold axis was extracted (Figure 3.6b) and its volume indicates a mass of ~8 kDa at 1σ. The similar trimeric density region at the local 3-fold axis was also extracted and its volume indicates a mass of ~12 kDa at 1σ. The larger volume corresponds well with that predicted for three copies of a domain identified in the N-terminal third of protein IX that is highly conserved among human and animal adenoviruses and includes residues crucial for incorporation of protein IX into the virion (98). Thus, we propose that the trimeric density corresponds to three copies of the conserved virus interaction domain of protein IX (aa 8-39).

Weak density was observed extending outward from the trimeric protein IX density (Figure 3.6c and 3.6d). This density is the correct length to correspond to an extended 7 residue polypeptide chain. Similar weak density was also observed above the local 3-fold trimeric density regions. We propose that the weak density corresponds to the N-terminal 7-residues of the Ad5 protein IX, which immediately precede the conserved virus interaction domain. We note that three hexon HVR2 loops are directly over the N-terminal protein IX density (Figure 3.6c and 3.6d). This may explain why the N-terminal domain of protein IX appears to be shielded from antibodies in immunoelectron microscopy studies (3).

It has been noted that the middle third of the protein IX sequence contains an alanine-rich region (aa 60 to 72) that might be flexible, and that the C-terminal third has a predicted coiled coil (98). We submitted the Ad5 protein IX sequence to the MultiCoil web server (142) and found that a dimeric coiled coil is predicted for 39 residues (aa 97-135) with a high probability (0.86 - 0.99). Thus we searched the cryoEM difference density for any region that might correspond to a coiled coil. Other cryoEM structures have demonstrated that at 9 Å resolution it is possible to resolve bundled α-helices (26). We found density that might correspond to a single 39-residue α-helix located between two peripentonal hexons (Figure 3.7a). No well-defined difference density was found that might correspond to a dimeric coiled coil of this length. The only other similar rod-shaped density in the difference map was near the 2-fold symmetry axis; however, it was not as well defined.

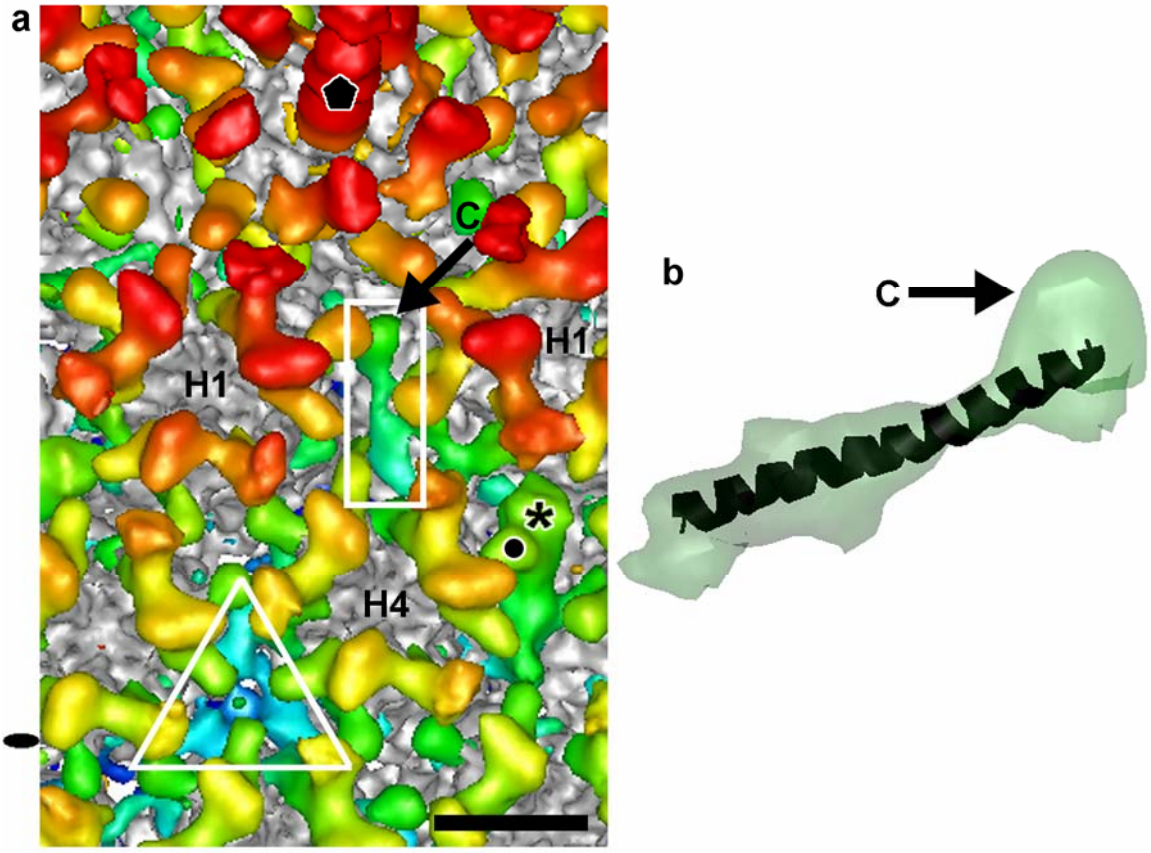


Figure 3.7. C-terminal domain of protein IX. **(a)** A region of the cryoEM difference density including a penton base, two peripentonal hexons (H1), one hexon in position four (H4), rod-shaped density assigned to the C-terminal predicted α-helix of protein IX (green, outlined with a white rectangle), and nearby protein IX N-terminal density (cyan, outlined with a white triangle). The asterisk marks the density assigned to a portion of polypeptide IIIa and the round dot indicates the interacting HVR4 loop from a neighboring hexon. Crystallographic symbols are shown as in Figure 3.6a. Bar, 50 Å. **(b)** A side view of the rod-shaped density (transparent green) fitted with a model for a 39-residue α-helix (black) as predicted for protein IX residues 97-135. Arrows mark the teardrop-shaped density lobe assigned to the C-terminal 5-residues (aa 136-140) following the predicted α-helix. Figure reprinted from (104) with permission from Elsevier.

The rod-shaped density between the peripentonal hexons has a teardrop-shaped region at one end that is similar in shape and size, or just slightly smaller than the hexon HVR4 protrusion. We assigned the hexon HVR4 density lobe to a 6-residue gap in the hexon crystal structure. Thus, it is reasonable to assume that the teardrop-shaped region corresponds to approximately 5 or 6 residues. A model for a 39-residue α-helix fits nicely within the rod-shaped density (Figure 3.7b). No other capsid protein has a predicted α-helix of 39 residues or longer. Therefore, we propose that the rod-shaped density corresponds to a single copy of the protein IX predicted α-helix (aa 97-135), and that the teardrop-shaped density at the end corresponds to the C-terminal 5 residues (aa 136-140). There is some weak difference density that might correspond to the flexible midsection of protein IX spanning half of gap between the assigned N- and C-terminal portions of protein IX.

Discussion

Here we present a subnanometer resolution structure of the Ad35F vector. This vector is ~100 fold more efficient at gene delivery to human monocytic cells than the Ad5F vector, which differs only in its displayed fiber protein. The fibers of the various adenovirus serotypes vary in length and flexibility, which has been demonstrated to affect attachment receptor usage and cell tropism (115, 144). The cryoEM reconstruction shows that the Ad35 fiber is flexible, as predicted by sequence analysis. In particular repeat 3 of the Ad35 fiber shaft has an insert of four residues that likely disrupts the β-spiral fold. It has been noted that the third repeat in most adenovirus serotypes, except those of subgroup D, has a nonconsensus sequence and is likely to induce flexibility in the fiber

shaft (145). We have previously reconstructed the full length of a subgroup D (Ad37) fiber, which indicates that it is rigid and straight (23). Both Ad35 (subgroup B) and Ad37 (subgroup D) utilize CD46 as the attachment receptor. However, Ad37, but not subgroup B Ads, has a strict requirement for calcium ions for efficient CD46 interaction (109, 146). Together with the observed difference in flexibility between the Ad35 and Ad37 fiber shafts noted in this study, these findings suggest that subgroup B and D adenoviruses have differences in receptor engagement. Our recent studies show that both viruses require the CD46 short consensus repeats SCR1 and SCR2, but they could bind to different sites within these domains and with different affinities (L. Pache and G.R. Nemerow, unpublished results).

By positioning the hexon and penton base crystal structures within the 9 Å resolution cryoEM density of Ad35F, a pseudo-atomic model for a capsid of 240 hexon trimers and 12 penton bases was generated. This calculated capsid was subtracted from the cryoEM density to produce a difference map that revealed density from numerous surface loops of hexon and penton base. In comparing the cryoEM and x-ray structures, we used the atomic models of hexon and penton base, rather than the crystallographic electron density maps. If we had used the crystallographic electron density maps, we suspect that the cryoEM and crystallographic structures would have appeared more similar. We analyzed the difference density on top of the hexon in terms of the gaps in the crystal structure and found that these density regions correspond to several antigenic hypervariable regions. In addition, this high resolution cryoEM analysis enabled us to predict with higher accuracy the epitope region of the 9C12 neutralizing monoclonal antibody, studied earlier at only moderate (21 Å) resolution (134). Remarkably,

difference density was observed for numerous sidechains over the surface of hexon, including those of the hypervariable HVR2 loop that were disordered in the crystal structure. Difference density was also found for the integrin-binding RGD loop and the variable loop on the side of the penton base. The variable loop is actually present in the crystal structure and thus our result indicates that it is flexible and occupies a larger volume than would be expected for a rigid loop. This loop has been proposed as a possible site for vector retargeting insertions (36). This cryoEM study suggests that insertions could be tolerated at this point within the adenovirus capsid, especially if a flexible linker region were added.

The cryoEM difference map should include the ordered portions of the minor capsid proteins. The earlier moderate resolution cryoEM difference map analysis suggested that protein IIIa, or at least part of IIIa, is on the outer capsid surface along the edges of the icosahedral facet (123). A similar density lobe is found on the outer capsid surface in the 9 Å resolution difference map, which could account for approximately half of the mature form of protein IIIa. In addition, we have found internal density associated with hexons and with the penton base. The density below the penton base might correspond to protein IIIa, VI, or VIII or perhaps to the adenovirus protease. The copy number is correct for protein IIIa, with 60 copies per virion (61). Volume analysis indicates that it might be one domain of the 63 kDa mature form of protein IIIa. This location for protein IIIa would be consistent with the finding that penton base and protein IIIa are the most thermally labile components of the virus (96). It would also be consistent with the Argon plasma etching result indicating that protein IIIa is located between the capsid and the core (82). Additional information will be needed to position

protein IIIa within the capsid. We speculate that the internal density associated with hexon might correspond to protein VI, since it approaches the hexon residues implicated in protein VI binding (72, 141). If the minor capsid components are partially or fully disordered and do not follow the icosahedral symmetry of the virion, they will be difficult to definitively locate.

Immunoelectron microscopy indicates that the C-terminus of protein IX, but not the N-terminus, is accessible on the viral surface (3). In addition, several groups have successfully attached domains to the C-terminus of protein IX and incorporated these fusion proteins into adenovirus vectors (90). The cryoEM structural analysis presented here has led to a more complete model for the location of protein IX within the adenovirus capsid (Figure 3.8). Four trimeric regions of cryoEM density were observed within an icosahedral facet in the sites traditionally thought to be the protein IX binding sites. Volume analysis suggested that this density corresponds only to the N-terminal conserved domain of protein IX known to be crucial for capsid association. The cryoEM density also reveals weak density above this domain that is the right size for the N-terminal, presumably flexible, 7 residues. The locations we have proposed for all twelve copies of the N-termini of protein IX within the icosahedral facet are shielded by hexon HVR2 lobes and thus would not be accessible to antibodies, in agreement with the immunoelectron microscopy results.

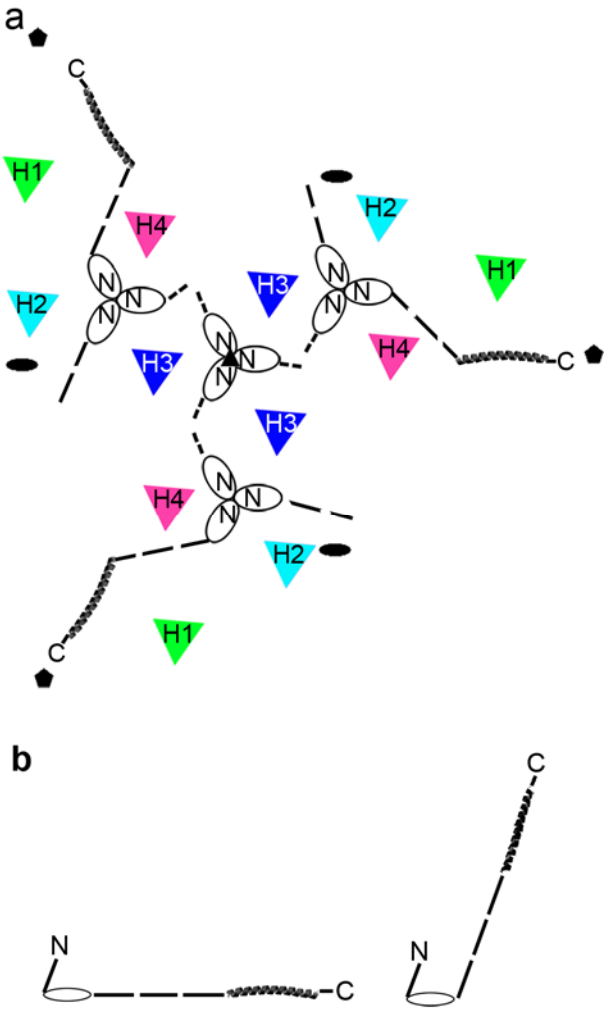


Figure 3.8. Diagram of protein IX within an icosahedral facet. **(a)** The positions of N- and C-termini are indicated; the conserved N-terminal region (aa 8-39) is symbolized by an oval outline; the flexible midsection (aa 44-96) by a dashed line; and the predicted C-terminal α-helix (aa 97-135) by a helix. Three copies of the C-terminal α-helices are proposed to bind on the capsid surface near the icosahedral 5-fold axes. Well-defined density is not observed for the other C-terminal domains and these may extend outward from the capsid surface. The four unique hexon positions within the asymmetric unit are labeled H1 through H4. Crystallographic symbols are shown as in Figure 3.6a. **(b)** Side view of the two proposed protein IX conformations: with the C-terminal α-helix binding on the capsid surface (left), and with the C-terminal α-helix extending outward from the capsid surface (right). Figure reprinted from (104) with permission from Elsevier.

The cryoEM density suggests that three of the twelve copies of protein IX within one facet have their C-terminal domains binding on the capsid surface. We found well-defined rod-shaped density between peripentonal hexons that agrees with a single 39-residue α-helix with a loop of 5 residues at the extreme C-terminus. There is no well-defined difference density that can be easily assigned to the C-terminal domains of the other nine copies of protein IX within one facet. Therefore, we propose that the flexible alanine-rich midsection may allow protein IX to adopt two different conformations (Figure 3.8b). Perhaps some of the protein IX C-terminal domains bind to the capsid and others extend away from the capsid surface and are not reconstructed. It is possible that pairs of extended C-terminal domains form coiled coils above the capsid surface. An extended conformation for some copies of protein IX would make these C-termini accessible to antibodies and provide plenty of space for large C-terminal additions.

The 9 Å resolution cryoEM structure of adenovirus provides a wealth of new knowledge for this human pathogen, which is also recognized as important for potential gene therapy and vector-based vaccine applications. In particular this structural investigation has provided insight into the flexible regions of the three major capsid proteins, hexon, penton base and fiber. In addition, protein IX is proposed to have some flexibility and this property may be advantageous for its use as a platform for ligand additions. This study has also shown that not all of the adenovirus capsid components follow strict icosahedral symmetry. Perhaps flexibility in the adenovirus capsid is essential for proper assembly and maturation, in addition to receptor engagement during cell entry.

CHAPTER IV

VISUALIZATION OF α-HELICES IN A 6 Å RESOLUTION CRYO-EM STRUCTURE OF ADENOVIRUS ALLOWS REFINEMENT OF CAPSID PROTEIN ASSIGNMENTS

Aim

Three points were unaddressed following the cryoEM structural analysis of Aim I: (1) the location of the predicted coiled coil of protein IX; (2) the location of half of the protein IIIa density; and (3) the identity of the density associated with the penton base along the inner capsid surface. To help address these questions, the objective of Aim II was to solve a high resolution structure of Ad35F in order to visualize α-helices as density rods. We reasoned that if α-helices were resolved as rods in capsid proteins with known atomic structures such as penton base and hexon, then α-helices could be identified in other capsid density. We hoped to use this information in conjunction with α-helical predictions for proteins IIIa, VI, VIII, and IX to interpret capsid density with greater confidence than was possible at lower resolutions. This chapter is an expanded version of a paper in press in the Journal of Virology (105).

Experimental design

The resolution of the first Ad35F structure leveled-off at 9 Å, so additional datasets were necessary to refine the structure to higher resolution. In order to increase the number of images in the reconstruction by approximately four-fold, dataset 2 was collected with the Leginon semiautomated data acquisition software package (124). The 9

Å resolution structure of Ad35F was calculated from images collected with a nominal magnification of 200,000 and a pixel size of 0.6 Å at the specimen level. Since this data was collected on a CCD camera, the most efficient data transfer occurred between low resolution and ~9 Å, which could be a factor limiting the resolution of the first structure of Ad35F. Shifting to a higher magnification like 310,000 (0.38 Å/pixel) extends the efficient information transfer region to ~6 Å. Thus, dataset 3 was collected at 310,000X to amplify 6 Å information in the reconstruction.

For the 9 Å resolution structure of Ad35F, the most astigmatic images were excluded from processing and the three-dimensional structure. The astigmatism in the remaining images, though less in magnitude, was simply ignored. This could be another factor limiting the resolution, so we employed a second image processing program, Frealign (Fourier Reconstruction and Alignment), which takes the astigmatism of each image into account when correcting for the contrast transfer function (CTF) imposed by the microscope.

Since the goal of Aim II is to resolve α-helices, we strove to extend the resolution of the Ad35F reconstruction to the range of 5 – 7 Å. By the Nyquist Theorem, 5 Å resolution is possible if the pixel size is at most 2.5 Å; however, resolution better than three times the pixel size is rarely achieved. In our case when we binned the particle images to 500^2 pixels, three times the pixel size (2.4 A) is 7.2 Å. When the resolution of the reconstruction reached ~7 Å, dataset 3 was rebinned from the original micrographs to 750^2 pixels and processed with a finer pixel size (1.5 Å) in order to facilitate extension of the resolution out to possibly 4.5 Å. The size of the reconstructed volume was then 750^3 voxels, which necessitated using ~20 GB of memory (RAM) for image processing with

Frealign. This extremely large memory requirement dictated that the final processing had to be performed on an SGI Altix computer equipped with 32 GB of RAM.

CryoEM structure at 6Å resolution reveals α-helices

The cryoEM structure of Ad35F at 9 Å resolution provided new information about the capsid proteins, particularly for protein IX, but a complete assignment was not possible for the additional capsid proteins (104). Acquisition of more particle images (6,880 *vs*. 2,215 total), inclusion of more particle images in the final reconstruction (2,630 *vs*. 964), and improvement in the image processing methods has enabled us to extend the resolution and enhance our interpretation of adenovirus structure (Figure 4.1). The resolution of the new Ad35F reconstruction as assessed by the conservative 0.5 Fourier shell correlation (FSC) threshold is 6.9 Å (Figure 4.1b). Information extends to 6.1 Å as indicated by the 0.3 FSC threshold.

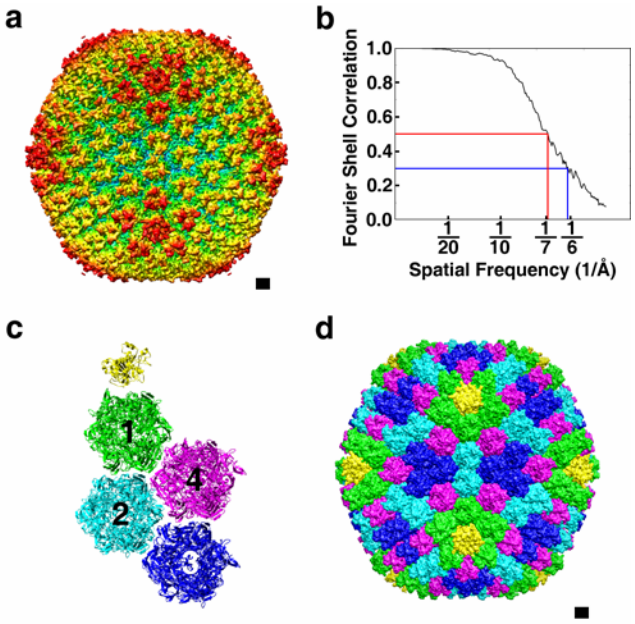


Figure 4.1. CryoEM structure of Ad35F at 6 Å resolution. **(a)** Surface view along a 2-fold symmetry axis of the reconstruction. The density is radially color-coded (red = 596Å; blue = 316Å). The flexible fibers were masked during image processing and only a short portion of the fiber is reconstructed. **(b)** A Fourier shell correlation (FSC) plot indicating a resolution in the range of 6.1 to 6.9 Å, at the 0.3 and 0.5 thresholds respectively. **(c)** Hexon (PDB 1P30) (103) and penton base coordinates (PDB 1X9T) (157) for one asymmetric unit of the capsid after docking within the cryoEM density. The penton base monomer is in yellow, hexon in position 1 in green, position 2 in cyan, position 3 in blue, and position 4 in magenta. **(d)** Pseudoatomic capsid calculated by applying icosahedral symmetry to the asymmetric unit. Scale bars, 50 Å. Figure reprinted from (105) with permission from the American Society for Microbiology.

The 6 Å resolution map enabled a more accurate fitting of the atomic resolution structures of hexon (103) and penton base (157) within the cryoEM density with the CoLoRes quantitative docking tool in the Situs software package (15). In particular the magnification was refined to within +/-0.25% by fitting with the hexon crystal structure, whereas at 9 Å resolution the magnification could only be refined to within +/-2.0% by this method. After docking the hexon and penton base crystal structures into the cryoEM density, we generated coordinates for the asymmetric unit of the adenovirus capsid including four independent hexon trimers and one monomer of the penton base (Figure 4.1c). Application of icosahedral symmetry to the asymmetric unit produced a pseudoatomic model for the adenovirus capsid containing all 240 copies of the hexon trimer and 12 copies of the pentameric penton base, representing over 5.2 million non-hydrogen atoms (Figure 4.1d).

The 6 Å resolution reconstruction is significant in that α-helices longer than 10 amino acids are now resolved within the icosahedral capsid. The backbone of a 10 residue α-helix has nearly three complete helical turns, and this is long enough to produce a rod-like shape that is recognizable at 6 Å resolution (55). Surface representations of the major capsid proteins penton base and hexon also reveal finer detail than observed in the 9 Å resolution map (Figures 4.2a and 4.3a). The Ad35 fiber was previously noted to be flexible (104), and therefore it was masked during image reconstruction. Thus, only a short portion of the fiber shaft is visible protruding from the penton base in the reconstruction (Figure 4.2a). The crystal structure of penton base (157) indicates that the monomer has two α-helices longer than 10 amino acids, corresponding to 10 long α-

helices in the pentamer (Table 4.1, Figure 4.2b). These penton base α-helices are clearly resolved as density rods in the cryoEM reconstruction (Figure 4.2c, d, e).

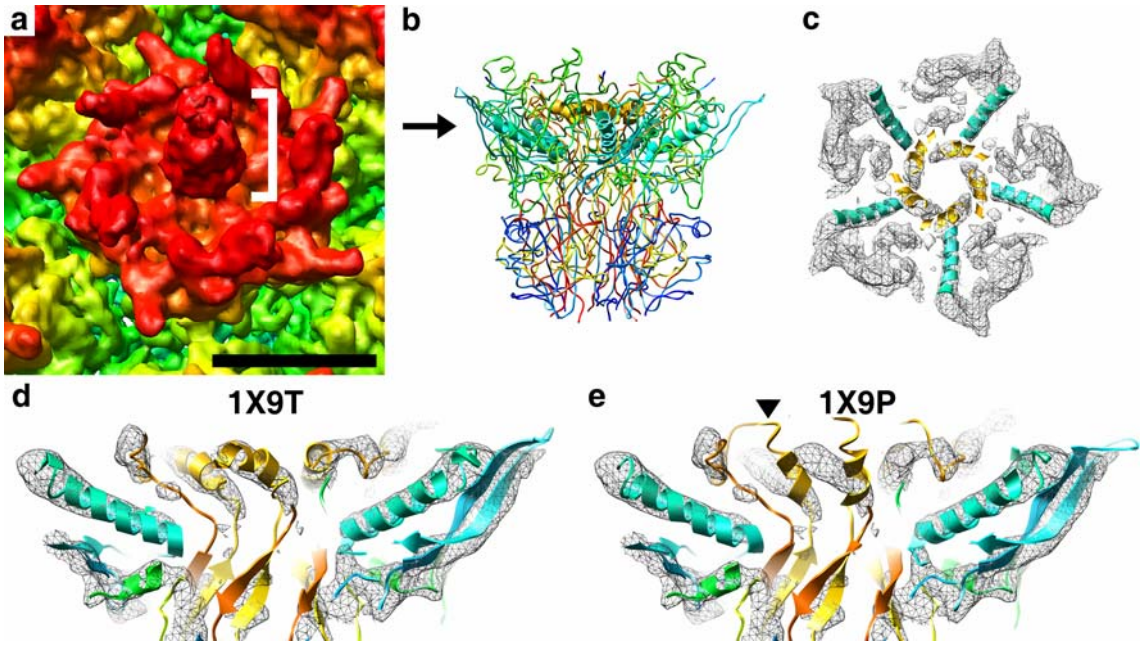


Figure 4.2. Penton base α-helices resolved in the cryoEM reconstruction. **(a)** Surface representation of the penton base with a short portion of the protruding fiber indicated by a white bracket. Scale bar, 50 Å. **(b)** Penton base coordinates (PDB 1X9T) (157) with the α-helices of 10 or more residues shown as helical ribbons. The ribbon structure is rainbow colored with the N-terminus in blue and the C-terminus in red. The arrow indicates the position of the plane in (c). **(c)** Crop plane through penton base density showing that the two long α-helices per monomer (PDB 1X9T) are resolved as rods within the cryoEM density (mesh). **(d)** A perpendicular crop plane through penton base showing the agreement of the cryoEM density with the α-helices in the 1X9T penton base crystal structure (with a fiber peptide). **(e)** The same crop plane as in (d) but with the cryoEM density aligned with the 1X9P penton base crystal structure (without fiber). The arrowhead indicates the conformational switch region affected by fiber binding for one penton base monomer, which includes the α-helix (aa 482-491) shown in gold. Figure reprinted from (105) with permission from the American Society for Microbiology.

Table 4.1. Long α-helices* in the penton base crystal structure (PDB ID 1X9T) (157)
*Only α-helices of 10 or more residues are listed.

Helix	Residues	Length (aa)
h1	176 - 193	18
h2	482 - 491	10

Two alternative crystal structures of the penton base have been solved: one in complex with an N-terminal fiber peptide (PDB 1X9T) and another without fiber (PDB 1X9P) (157). Fiber binding drives a conformational switch in penton base residues 482-505, which includes one of the two moderately long α-helices. The cryoEM density of penton base is clearly more consistent with the conformation of 1X9T (with fiber peptide) than that of 1X9P (without fiber) (Figure 4.2d,e), which is as expected since the Ad35F vector includes the Ad35 fiber. This result demonstrates the level of accuracy with which the 6 Å resolution cryoEM structure of Ad35F can be interpreted.

Surface representations of hexon in the 6 Å resolution cryoEM reconstruction (Figure 4.3a) show a sharp protrusion for the HVR4 hypervariable region (103). The hexon crystal structure is missing six residues (aa 251-256) in the HVR4 region due to disorder (103). The cryoEM density suggests that these residues form an elongated loop protruding from the side of the hexon. The hexon PDB coordinate file was submitted to the CODA protein loop prediction web server (30) and several reasonable model loop structures were returned. The top ranked model generated by CODA fit best within the cryoEM density (Figure 4.3b). The crystal structure of hexon indicates that each monomer has six α-helices longer than 10 amino acids, corresponding to 18 long α-helices in the trimer (Table 4.2, Figure 4.3c). The agreement between the cryoEM density rods and the α-helices in the crystal structure is shown for a hexon in position 4 of the asymmetric unit (Figure 4.3d). Two of these α-helices (h1 and h5) are at the bottom of the trimer. In the 9 Å resolution structure of Ad35F (104), these two helices were not fully resolved as cylinders but rather blended in with the surrounding non-helical density (Figure 4.3e). In addition all of these α-helices in the 6 Å resolution structure are resolved

as rods within each of the three other independent hexons in the asymmetric unit (data not shown). These results for hexon and penton base demonstrate that the resolution achieved for this cryoEM structure is sufficient to reliably identify α-helices within the adenovirus capsid. They also provide confidence for interpretation of the α-helices observed in the remaining cryoEM density for the additional capsid proteins.

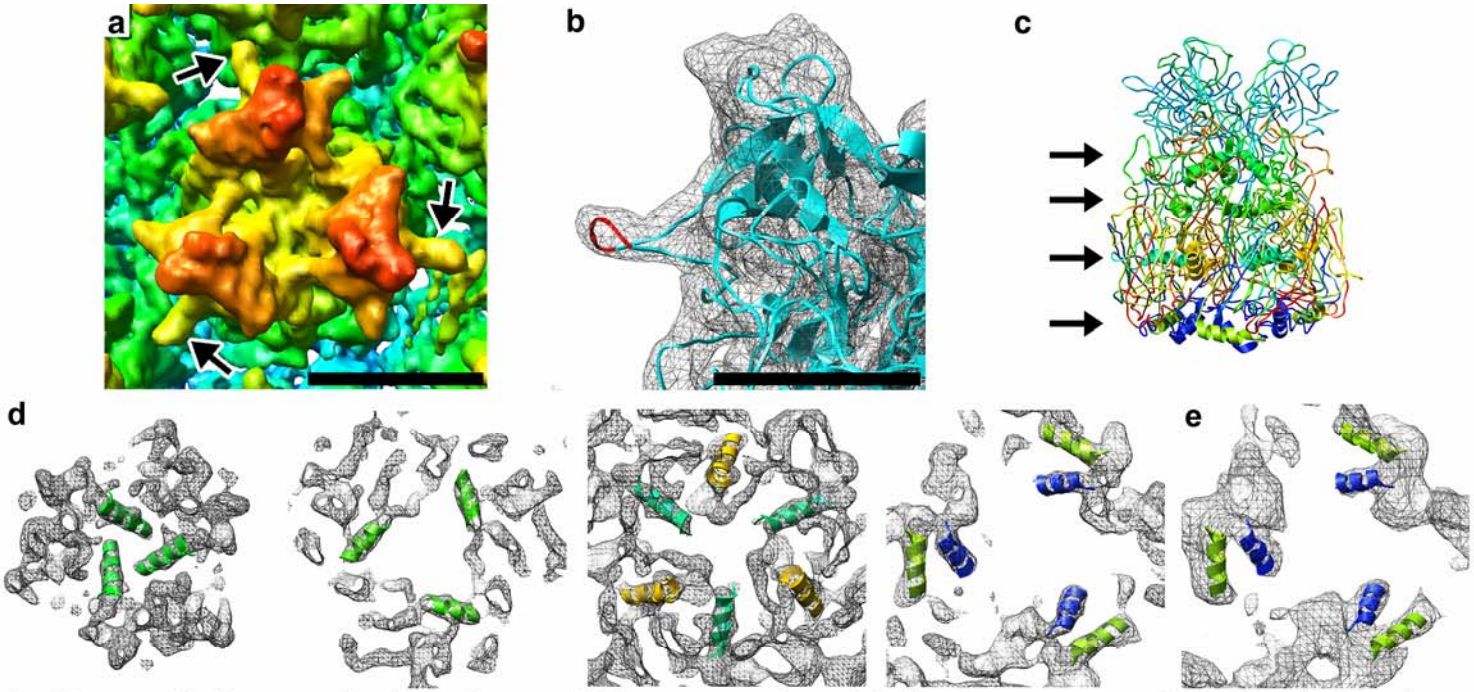


Figure 4.3. Hexon α-helices resolved in the cryoEM reconstruction. **(a)** Surface representation of a hexon, in position 4 of the asymmetric unit, with the sharp protrusions of the HVR4 region (aa 251-256) indicated by arrows. Scale bar, 50 Å. **(b)** Enlarged side view of the hexon tower density (mesh) with the docked trimer coordinates (PDB 1P30) (cyan) and a model for loop residues (aa 251-256) (red). Scale bar, 25 Å. **(c)** Hexon trimer coordinates with the α-helices of 10 or more residues shown as helical ribbons. The ribbon structure is rainbow colored with the N-terminus in blue and the C-terminus in red. The arrows indicate the positions of the four planes in (d). **(d)** Crop planes through a position 4 hexon showing that all of the moderately long α-helices in the hexon trimer are resolved as rods within the cryoEM density (mesh). **(e)** Crop plane from the 9 Å resolution cryoEM reconstruction similar to the rightmost panel in (d). Helices h1 (blue) and h5 (green) that are resolved at 6 Å resolution (d,right) are not quite resolved from nearby density at 9 Å resolution (e). Figure reprinted from (105) with permission from the American Society for Microbiology.

Table 4.2. Long α-helices* in the hexon crystal structure (PDB ID 1P30) (103)
*Only α-helices of 10 or more residues are listed.

Helix	Residues	Length (aa)
h1	24 - 35	12
h2	364 - 377	14
h3	462 - 476	15
h4	539 - 550	12
h5	622 - 634	13
h6	761 - 774	14

Secondary structure predictions for the additional capsid proteins

In addition to hexon, penton base, and fiber, the adenovirus capsid also contains proteins IIIa, VI, VIII, and IX originally named for their mobility in polyacrylamide gels. Currently no atomic resolution structural information is available for these four proteins. Difference mapping, or subtraction of the pseudoatomic capsid of hexon and penton base from the cryoEM reconstruction, should reveal density for the additional capsid proteins if they are packaged with icosahedral symmetry. The adenovirus core, composed of the viral DNA as well as proteins V, VII, mu, and terminal protein, does not follow the icosahedral symmetry of the capsid and thus is not well reconstructed. The observation of α-helices within the cryoEM reconstruction opens up the possibility of assigning density to the additional capsid proteins based on their predicted α-helical content, as well as their known copy numbers and expected capsid locations.

PsiPred secondary structure prediction (12) was performed for Ad5 proteins IIIa, VI, VIII, and IX. Only predicted α-helices of 10 or more residues are reported here, since these are the α-helices that should resolve at 6 Å resolution. Protein IIIa is likely to be highly α-helical, with 14 α-helices in the N-terminal two thirds of the sequence (aa 1-400) and 2 α-helices in the C-terminal third (aa 401-570) (Table 4.3). The first half of the mature form of protein VI is also likely to be α-helical, with two predicted α-helices of 20 and 46 residues (Table 4.4). Submitting the full-length sequence of the precursor to protein VIII to the PsiPred server indicates one long α-helix of 24 residues in the 12.1 kDa fragment (aa 1-111 of the precursor) and no long α-helices in the 7.6 kDa fragment (aa 158-227 of the precursor) (Table 4.5). For protein IX, the PsiPred server indicates no α-helices for the viral interaction domain (aa 1-39), two possible α-helices of 10 and 20

residues in the alanine-rich low complexity region (aa 60-90) in the middle of the sequence, and one long α-helix of 40 residues (aa 92-131) in the C-terminal half of the sequence (Table 4.6). The COILS, PAIRCOILS, and MULTICOILS web servers (4, 68, 142) all indicate a strong propensity for coiled coil formation for the C-terminal α-helix of protein IX. In contrast, the coiled coil web servers indicate that coiled coils in proteins IIIa, VI, and VIII are unlikely.

Table 4.3. Predicted α-helices* in Protein IIIa
*Secondary structure prediction was performed using the PsiPred Server (12). Only predicted α-helices 10 residues or longer are listed.

Helix	Residues	Length (aa)
h1	25 – 37	13
h2	50 – 61	12
h3	70 – 82	13
h4	88 – 105	18
h5	110 – 132	23
h6	137 – 149	13
h7	161 – 173	13
h8	230 – 242	13
h9	253 – 264	12
h10	272 – 285	14
h11	290 – 301	12
h12	316 – 335	20
h13	340 – 352	13
h14	360 – 376	17
h15	512 – 524	13
h16	527 – 536	10

Table 4.4. Predicted α-helices* in Protein VI
*Secondary structure prediction was performed using the PsiPred Server (12). Only predicted α-helices 10 residues or longer are listed.

Helix	Residues	Length (aa)
h1	39 – 58	20
h2	62 – 107	46

Table 4.5. Predicted α-helices* in Protein VIII
*Secondary structure prediction was performed using the PsiPred Server (12). Only predicted α-helices 10 residues or longer are listed.

Helix	Residues	Length (aa)
h1	38 – 61	24

Table 4.6. Predicted α-helices* in Protein IX
*Secondary structure prediction was performed using the PsiPred Server (12). Only predicted α-helices 10 residues or longer are listed.
[1]Helix h3 is predicted to form a coiled coil.

Helix	Residues	Length (aa)
h1	57 – 76	20
h2	79 – 88	10
h3[1]	92 – 131	40

Reassignment of protein density on the capsid exterior

In the 9 Å resolution difference map from Aim I, four trimeric regions of density per icosahedral facet and a large density lobe near each facet edge on the exterior of the capsid were observed (104). The trimeric density regions were confidently assigned to protein IX, since these sites had previously been shown by STEM to be occupied by protein IX (40). More precisely, volume analysis indicated that the trimeric regions corresponded to only the N-terminal virus interaction domains of protein IX. The density lobes observed at the facet edges were tentatively assigned as a domain of protein IIIa, in accord with an earlier cryoEM interpretation made at ~25 Å resolution (123). This assignment for protein IIIa was based mainly of the fact that the copy number of these lobes was in agreement with that of protein IIIa, with 60 copies per virion (61, 123), and also on an early observation that a least a portion of protein IIIa might be exposed on the outside of the capsid (34).

The 6 Å resolution difference map generated by subtracting the pseudoatomic capsid from the cryoEM density also shows the four trimeric density regions in the middle of the facet and density lobes at the facet edges (Figure 4.4a). The trimeric density regions now display more structural detail. There is no indication for α-helices within these regions, which is consistent with the PsiPred secondary structure prediction for the N-terminal region of protein IX. Therefore, we interpret each trimeric region to be three copies of the N-terminal region of protein IX possibly up to residue 56, which is just before the first predicted α-helix.

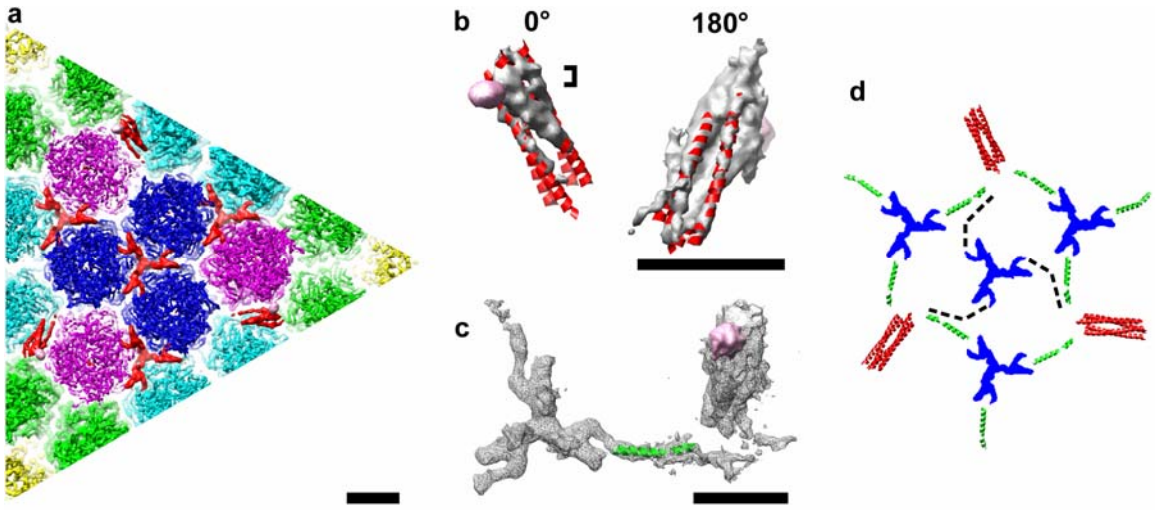


Figure 4.4. Protein density on the exterior of the capsid. **(a)** A region of the cryoEM difference map (red) and pseudoatomic capsid (colored as in Figure 1c) roughly corresponding to one icosahedral facet. Four trimeric density regions and three helical bundles are observed. Density from missing residues in the hexon and penton base crystal structures has been removed except for the hexon HVR4 loops (pink) that connect to the helical bundles. **(b)** Enlarged views of the helical bundle shown in two views related by a 180° rotation. The density is displayed with a lower isosurface in the 180° view to show the weak density corresponding to the fourth α-helix. The connected hexon HVR4 density is in pink. Four α-helices, each 40 residues long (red), are shown fit within the density. The bracket indicates the 5 Å ladder-like spacing corresponding to the expected sidechain spacing for successive α-helical turns. **(c)** Enlarged view of one trimeric density region and a nearby helical bundle shown with a low isosurface to reveal that they appear to be almost connected. Two α-helices of 10 and 20 residues (green) are shown fit within the connecting density. **(d)** Diagram of our interpretation of the difference map density on the exterior of the capsid. The four trimeric density regions are assigned to four trimers of the N-terminal capsid interaction domain of protein IX (aa 1-56) (blue); the weak density connections are assigned to the low complexity midsection of protein IX (aa 57-91) (green); and the helical bundles are assigned to the C-terminal predicted α-helix of protein IX (aa 92-131) (red). The black dashed lines indicate presumably extended (non α-helical) connections between the central trimeric density region and the helical bundles. Note that one of the four helices in each bundle appears to come from a copy of protein IX in an adjacent facet. Scale bars, 50 Å. Figure reprinted from (105) with permission from the American Society for Microbiology.

One of the most dramatic differences between the 6 and 9 Å resolution cryoEM structures is that at higher resolution the density lobe at the facet edge is resolved into a four-helix bundle (Figure 4.4b). The density lobe appears to be four moderately long α-helices twisting around each other as in a left-handed coiled coil. The cryoEM density clearly shows a 5 Å ladder-like spacing between the α-helices, which is consistent with the sidechain packing from sequential turns of an α-helix. The crystal structure of the tetrameric coiled coil domain of Sendai virus phosphoprotein (PDB 1EZJ) (126) served as good model for the observed coiled coil density. The 63 residue α-helix in the Sendai protein coiled coil was shortened to 40 residues, as predicted for the C-terminal protein IX α-helix, and manually fit within the cryoEM density (Figure 4.4b). The shortened four-helix bundle from the Sendai protein fit within the cryoEM density rods without any further modification. When displayed at a low isosurface value to include even weak density, the length of the adenovirus helical bundle appears to be ~40 residues long. Given that the secondary structural analysis indicates that of all the additional capsid proteins only protein IX has a strong propensity for coiled coil formation, the density at the facet edges is reassigned to the C-terminal region of protein IX.

The density for one α-helix in the adenovirus coiled coil is stronger than that of the other three α-helices. In fact the fourth α-helix on the back side of the coiled coil is quite weak. These observations may be explained by the fact that only one of the α-helices, that with the strongest density, appears to interact with a neighboring hexon. A clear connection is seen between the strongest α-helix and the HVR4 loop of a hexon in position four of the asymmetric unit (Figure 4.4b). Based upon the model for the hexon HVR4 loop (Figure 4.3b), HVR4 residues K251 and N254 are the most likely sidechains

to interact with the coiled coil, since the other sidechains in the model point in the wrong direction. The weak density for three of the four α-helices suggests partial occupancy for these sites. Thus, the proposal presented in Chapter III that the C-terminal domain of protein IX may adopt two conformations, either binding on the capsid surface or extending away from the capsid (104), is still valid. The model can be elaborated to state that the C-terminal α-helices of protein IX may either cluster into the helical bundles at the facet edges, or possibly sit at other low occupancy sites on the capsid surface, or extend away from the capsid surface. Both low occupancy sites and extended conformations would be difficult to reconstruct by cryoEM single particle methods.

As additional support for the new assignment for the C-terminal domain of protein IX, the trimeric density region closest to the helical bundle almost connects to the helical bundle at low isosurface values (Figure 4.4c). This density connection can be modeled as two α-helices of 10 and 20 residues, as predicted for the middle low-complexity region of protein IX. Other partial connection arms are observed in the difference map, including an arm from the trimeric density along the facet edge that extends to the adjacent facet. The distance between various trimeric density regions and the helical bundles varies, so the low complexity region may adopt different conformations to span different distances. No connections are observed between the trimeric density region in the center of the facets and the helical bundles; however, if these connecting peptide regions are in an extended (non α-helical) conformation they would not be observed at this resolution. It appears that at least one of the four α-helices in the bundle may come from a neighboring facet. Our interpretation of the protein density in the difference map on the exterior of the capsid surface is shown schematically

in Figure 4.4d. If all four sites of all 60 helical bundles in the capsid are occupied, that would account for 240 (4 x 60) C-terminal domains of protein IX, in agreement with the copy number for protein IX in the adenovirus capsid.

Internal capsid density assigned to protein VIII

The 10 Å resolution cryoEM analysis of Fabry *et al.* led to an assignment for protein VIII to two elongated regions of density per asymmetric unit on the inner capsid surface (36). We concur with this assignment, as the 6 Å resolution difference map also shows two elongated and structurally similar density regions in the asymmetric unit on the inner capsid surface (Figure 4.5); however, the two independent copies of protein VIII are difficult to discern in the Ad35F 6 Å difference map without density segmentation. All of the density on the inner capsid surface in the 6 Å difference map is shown in Figure 4.5a for one icosahedral facet. The facet includes three asymmetric units, so six copies of protein VIII should be visible. Three copies of protein VIII are well resolved and surround the three copies of hexon in position 3 at the center of the facet. The other three copies of protein VIII are partially obscured by large clusters of α-helices below the penton bases at the corners of the facet. When these α-helical clusters are computationally removed, the additional copies of protein VIII become apparent. The two independent copies of the protein VIII in the asymmetric unit are shown in Figure 4.5b. Each copy of protein VIII appears to interact with four hexons, although the hexons differ (2, 3, and 4; or 1, 2, and 4) depending on the protein VIII site. When the two independent volumes of protein VIII density are overlaid, they are nearly identical three-dimensional structures (Figure 4.5c). Examination of the protein VIII density indicates

only one long α-helix of ~24 residues (Figure 4.5d), as predicted by PsiPred for the 12.1 kDa fragment of protein VIII (Table 4.5).

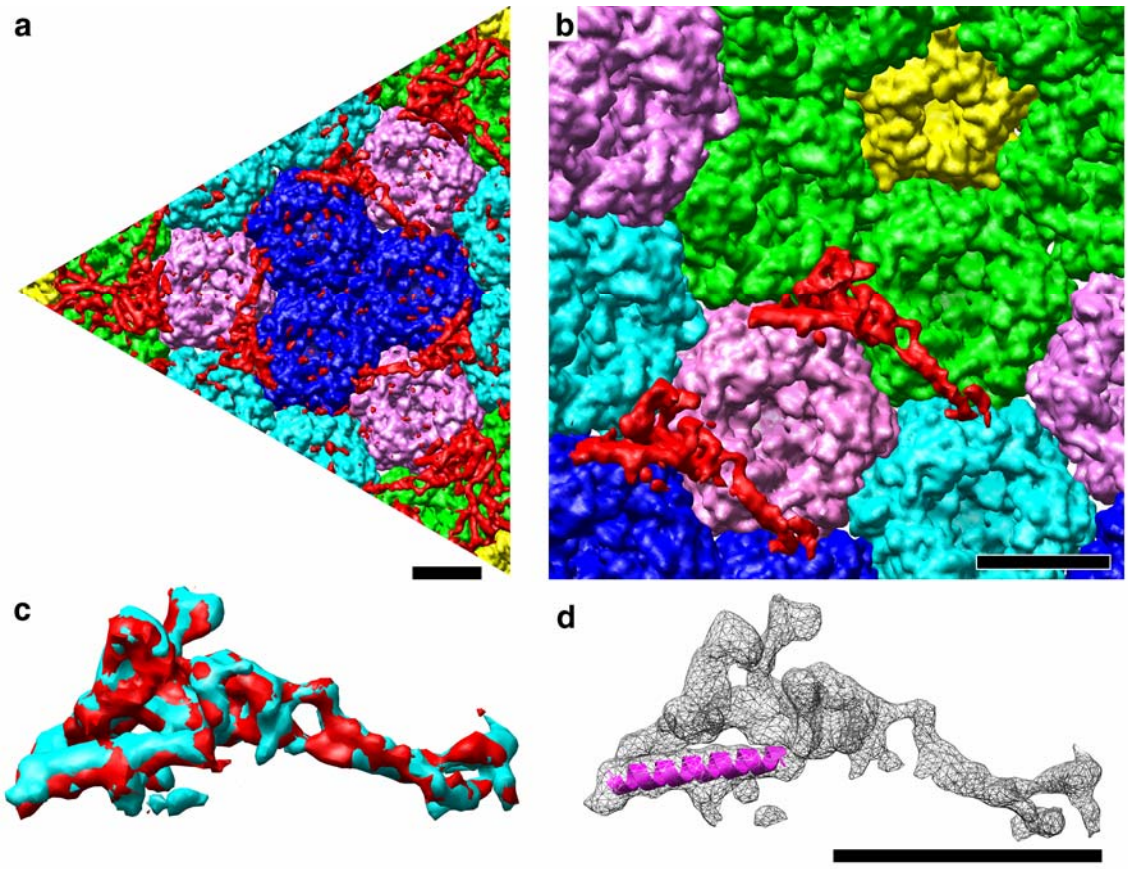


Figure 4.5. Protein density on the interior of the capsid including protein VIII. **(a)** A region of the cryoEM difference map (red) and pseudoatomic capsid (colored as in Figure 1c) roughly corresponding to one icosahedral facet viewed from inside of the capsid. Three separate triangular density regions previously assigned to protein VIII, as well as clusters of multiple α-helices below the penton base are observed. **(b)** Two independent copies of protein VIII (red) in the asymmetric unit. Density from other nearby proteins has been removed to reveal the similar three-dimensional structure of the two copies of protein VIII. **(c)** Overlap of the two independent copies of protein VIII density (red and cyan). **(d)** Fit of a 24-residue α-helix into the cryoEM density (mesh). Note that in this view other regions of protein VIII appear rod-shaped however interactive rotation of the density reveals that only one long α-helix is present as modeled. Scale bars, 50 Å. Figure reprinted from (105) with permission from the American Society for Microbiology.

Density assigned to protein VI within hexon cavities

The hexon trimer has a large central cavity that faces the viral core and residue 776, which has been implicated as a key residue in the hexon/protein VI interaction (141), is in a loop between two α-helices within the cavity (Figure 4.6). Thus, we carefully examined the difference map for any density within the hexon cavities. At a high isosurface value, only a few small, disconnected regions of density are observed in the hexon cavities and these disconnected regions appear as dots in Figure 4.5a. At a moderate isosurface value a disk-shaped region of density is observed at the top of each hexon cavity, as well as three rods of density near the bottom of the cavity (Figure 4.6). The disk-shaped density is surrounded by hydrophobic hexon residues at the top of the cavity including residues L463, L467, W468, and F471 from the three hexon monomers. The density rods come within 10 Å of the hexon loop with residue 776, which has been implicated in protein VI binding (141). At low isosurface the density disk at the top of the cavity connects to the three rods at the bottom of the cavity. Thus, all of the density within the hexon cavity is tentatively assigned to protein VI.

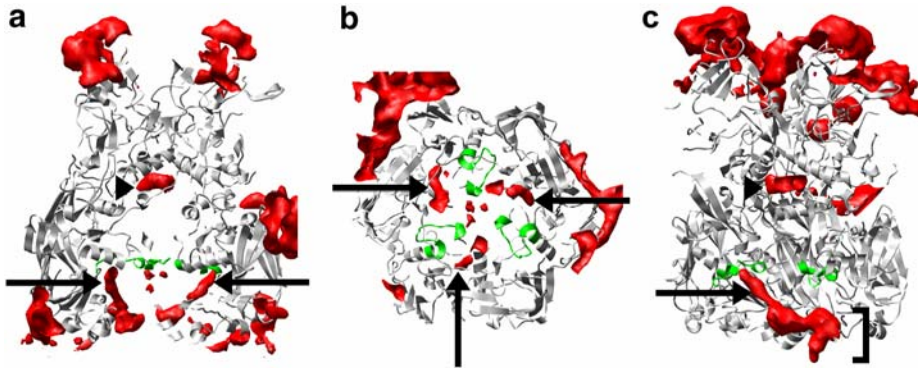


Figure 4.6. Density within the cavity of the hexon trimer tentatively assigned as protein VI. **(a)** Regions of the cryoEM difference map (red) near the hexon trimer in position 4 of the asymmetric unit. Only a thin plane of density through hexon is shown to reveal the central cavity. The hexon is oriented so that the exterior capsid surface is at the top and the interior surface is at the bottom. The hexon loop implicated in binding to protein VI is shown in green (aa 770-787). The density at the top of the hexon cavity is indicated by an arrowhead and two of the three density rods near the base of the cavity are marked with arrows. **(b)** A perpendicular thin plane of density showing the close proximity of hexon loop (aa 770-787) (green) to the rod-like density indicated by arrows. **(c)** Regions of the cryoEM difference map (red) near the hexon trimer in position 2 of the asymmetric unit. The plane is thicker than (a) and rotated in order to visualize the connection between the density rod (arrow) and the density folded over the base of the hexon (bracket). Figure reprinted from (105) and modified with permission from the American Society for Microbiology.

In contrast to the ~25 Å resolution structure of adenovirus that placed protein VI solely at the vertex region (123), the 6 Å structure of adenovirus reveals protein VI density within the cavities of all four hexons in the asymmetric unit. The density within the cavities also connects to density that folds over onto the bottom of the hexons, indicating that perhaps not all of the mature form of protein VI fits within the hexon cavity (Figure 4.6c). Similar density lining the hexon cavities was also observed at 9 Å resolution (104), but it is now resolving into subunits positioned on symmetry related sites on the inner surface of the hexon trimers. Consideration of stoichiometry indicates 720 possible binding sites for protein VI, or three potential binding sites in each of the 240 hexon trimers. However, mass spectrometry indicates only ~369 copies of protein VI per virion (61). Therefore, it seems likely that there is partial occupancy of all 720 possible binding sites for protein VI. This may explain the weak cryoEM density for protein VI in the Ad35F reconstruction.

A new assignment for protein IIIa on the inner capsid surface

A second major difference between the 6 and 9 Å resolution cryoEM structures is that the density below the penton base on the inner capsid surface is now resolved as large clusters of α-helices (Figure 4.7a). At 9 Å resolution this density was speculated to correspond to proteins IIIa, VI, or VIII or to the adenovirus protease (104). Now that α-helices are resolved at 6 Å resolution, a more informed assignment can be made. Protein IIIa is unique among the adenovirus capsid proteins in that secondary structure prediction indicates that it is a highly α-helical protein with an α-helical content of nearly 50%. PsiPred predicts at least 14 α-helices in the N-terminal two thirds and 2 α-helices in the

C-terminal one third of the sequence (Table 4.3). Each of the five helical clusters below the penton base contains at least 13 α-helices, in reasonable agreement with the predicted number for the N-terminal two thirds of protein IIIa. The copy number of this helical cluster is also consistent with mass spectrometry data for protein IIIa, with 60 copies of protein IIIa per virion, or one copy per asymmetric unit (61).

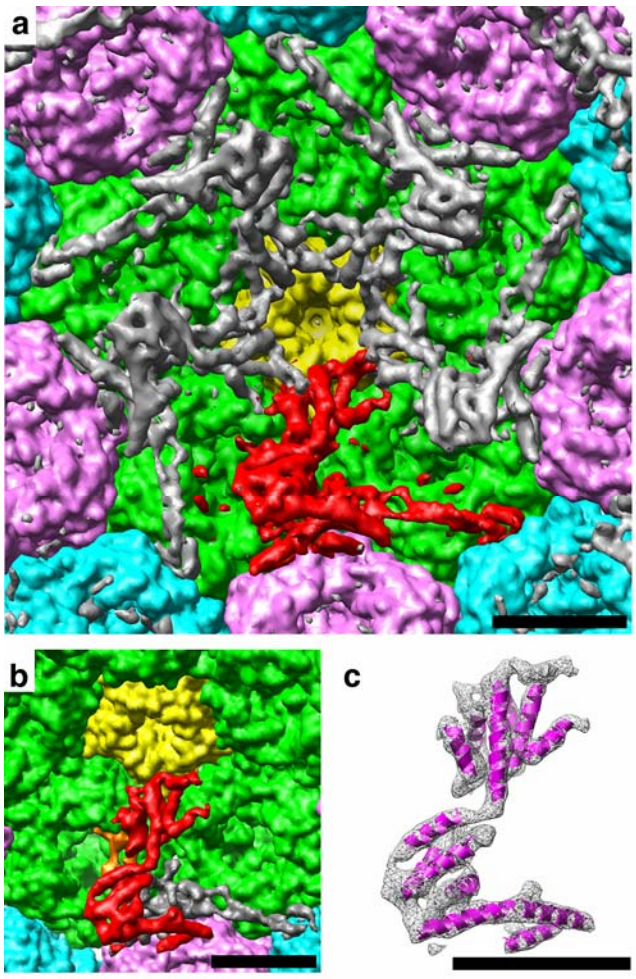


Figure 4.7. Internal capsid density assigned to protein IIIa. **(a)** Density in the difference map below the penton base includes protein VIII and clusters of multiple α-helices. One fifth of the density from the difference map is shown in red and the remainder in gray. A region of the pseudoatomic capsid is shown colored as in Figure 1c. **(b)** Enlarged view of one fifth of the density from the difference map. This includes one cluster of α-helices assigned to the N-terminal region (aa 1-400) of one protein IIIa monomer (red), together with one copy of protein VIII (gray), and unidentified density, possibly protein VI, at the base of a hexon (orange). **(c)** Fit of 13 α-helices, ranging in length from 12 to 23 residues, into the α-helical density assigned to protein IIIa (mesh). Scale bars, 50 Å. Figure reprinted from (105) with permission from the American Society for Microbiology.

Now that α-helices are resolved in the cryoEM density, none of the other alternative protein candidates for the density below the penton base seems reasonable. The crystal structure of the adenovirus protease (73) clearly does not fit within this α-helical density. Protein VIII is now accounted for in the difference map. Even if the current protein VIII assignment is incorrect, two copies of protein VIII per asymmetric unit would contribute only two long α-helices per asymmetric unit, which is not enough to account for the observed helical cluster of 13 helices per asymmetric unit. Protein VI is a possible, but highly unlikely, candidate for the helical cluster. PsiPred predicts two long α-helices in protein VI (Table 4.4), and given that there are ~369 copies of protein VI per virion this would imply ~6 copies per asymmetric unit that could in principle account for 12 α-helices per asymmetric unit. However, the helical cluster appears much more like a folded protein domain than a multimer with 2 helices contributed from each of 6 copies of protein VI. Also assigning the helical cluster to a multimer of protein VI is not very compatible with the ts147 mutant data indicating that protein VI binds in the hexon cavity (141).

Therefore, the most likely assignment for the helical cluster below the penton base is the N-terminal domain of protein IIIa, which has 14 predicted α-helices (Table 4.3). Before attempting to fit α-helices within the helical density, we segmented one helical cluster from its neighboring five-fold related clusters and from the nearby copy of protein VIII (Figure 4.7b). In addition, we segmented away a small region of density on the bottom of the hexon. Similar density is found on the bottom of all copies of hexon in the capsid, and thus it is likely to be protein VI or perhaps a core component that associates with the capsid.

After segmentation, 13 canonical α-helices ranging in length from 12 to 23 residues were fit within the helical cluster (Figure 4.7c). This accounts for all but one of the predicted α-helices in the N-terminal two thirds of protein IIIa. Clearly, at this resolution there is not a unique solution for placement of the α-helices within the density since in general the length of an α-helix can only be predicted to within a few residues and also since there are multiple predicted α-helices for protein IIIa with the same or similar length. Nevertheless the fit of α-helices presented in Figure 4.7c supports the assignment of the N-terminal region of protein IIIa to the helical cluster of density below the penton base.

Consideration of the 5-fold symmetry of the vertex region implies that 5 copies of protein IIIa bind in five symmetrically arranged areas on penton base as shown in Figure 4.7a. These 5 copies of protein IIIa nearly cover the entire accessible inner surface area of penton base. Based upon the pseudoatomic capsid, the penton base residues that are likely to interact with the helical cluster assigned to protein IIIa can be predicted (Table 4.7). The protein IIIa/penton base interface involves ~38 penton base residues, including residue 52, which is the first ordered residue in the crystal structure (Figure 4.8). It seems likely that the N-terminal tail of penton base (aa 1-51) is also involved in the interaction. The PsiPred server predicts just one α-helix with a high level of confidence for the N-terminal tail of the penton base, and it is only 7 residues long. At most one short helical density rod in the helical cluster assigned to protein IIIa might correspond to the N-terminal tail of penton base. This new assignment for protein IIIa below the penton base on the inner capsid surface is in accord with the biochemical findings that protein IIIa is

one of the more labile capsid components and is released shortly after the penton base in thermally- and low pH-induced disassembly assays (47, 96, 140).

Table 4.7. Penton base and hexon interactions with the protein IIIa density.
[1]Some of these residues are also within 10 Å of protein VI

Penton Base		**Hexon**	
Monomer 1	Monomer 2	Chain A	Chain B
52 – 66	116 – 117	5 – 24[1]	633 – 641
111 – 119	567 – 569	42 – 61[1]	669 – 673
532 – 534		89 – 91	877 – 878
564 – 569		620 – 639[1]	880 – 898[1]
		928 – 933	925 – 931
			938 – 945

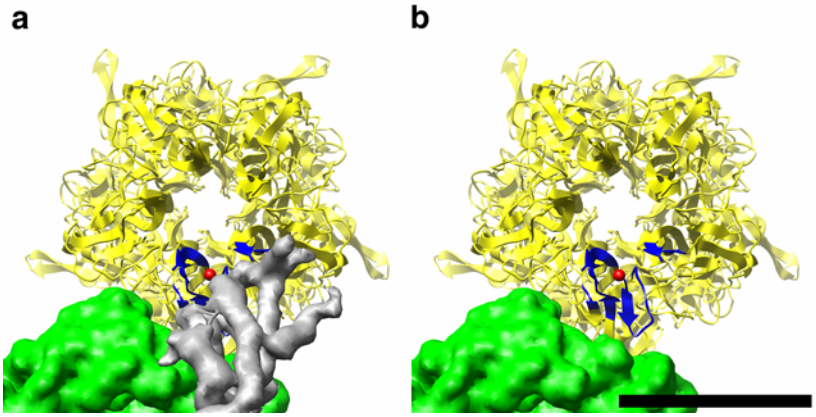


Figure 4.8. Protein interactions between protein IIIa and penton base. **(a)** A ribbon representation of penton base (157) (yellow) viewed along the 5-fold axis, and surface representations of portions of a neighboring hexon (green) and the density assigned to protein IIIa (gray). The residues in penton base that are within 10 Å of protein IIIa are in blue. **(b)** Same as (a) but with the protein IIIa density removed. The first residue (aa 52) resolved in the crystal structure of penton base is shown as a red sphere. Scale bar, 50 Å. Figure reprinted from (105) with permission from the American Society for Microbiology.

Protein IIIa forms an extensive interaction surface with a peripentonal (position 1) hexon (Figure 4.9). This interaction involves over 90 residues from two monomers within the trimer and may also include the first four N-terminal residues from one monomer that are missing from the crystal structure (Table 4.7). Although the interaction of the helical cluster is predominantly with one peripentonal hexon, portions of the helical cluster also interact with an adjacent peripentonal hexon and with a nearby hexon in position 4 (Figure 4.7a, b). Some of the hexon residues within 10 Å of protein IIIa are also within 10 Å of density likely to be protein VI (Table 4.7).

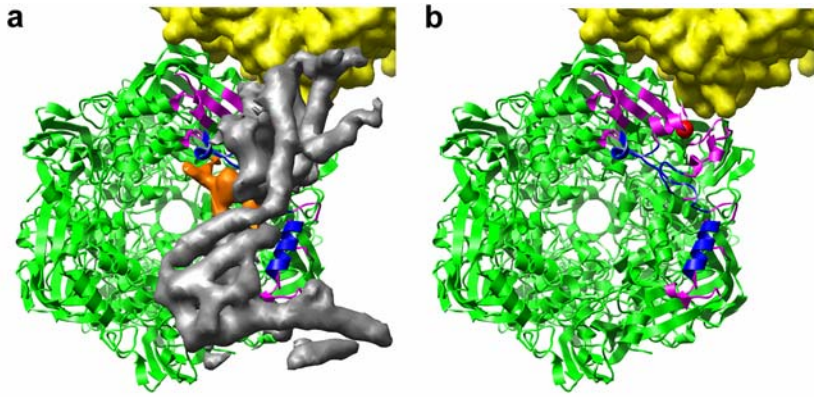


Figure 4.9. Protein IIIa interactions with hexon. **(a)** A ribbon representation of a peripentonal (position 1) hexon (103) (green) viewed along the trimeric 3-fold axis, and surface representations of a portion of penton base (yellow), the internal protein IIIa density (gray), and the density likely to be protein VI (orange). The residues in hexon that are within 10 Å of protein IIIa density are in magenta, and those with 10 Å of both proteins VI and protein IIIa are in cyan. **(b)** Same as (a) but with the protein IIIa and density likely to be protein VI removed to reveal the full interaction sites on hexon. The first residue (aa 5) resolved in the crystal structure of hexon monomer chain A is shown as a red sphere.

Discussion

The collection and averaging of more cryoEM particle images as well as the use of improved image processing techniques enabled us to determine a 6 Å resolution structure for Ad35F. This resolution is significant in that α-helices can now be resolved in the capsid. Docking of the hexon and penton base crystal structures allowed us to confirm that all α-helices of 10 or more residues are resolved as rods within the cryoEM density. One improvement that had a significant effect on the resolution was to process particle images with a smaller pixel size of 1.5 Å on the molecular scale. This increased the size of the reconstructed volume to 750^3, which in turn significantly increased the memory requirement to ~20 GB, but we found this to be essential in order to reach 6 Å resolution. A second improvement that led to higher resolution was correction for the astigmatism in the cryo-micrographs, which was accomplished with the Frealign software package (48).

Two significant new features were revealed in the 6 Å resolution adenovirus structure that were not apparent at 9 Å. First, a density lobe at the facet edge on the exterior of the capsid that had long been assumed to be a domain of protein IIIa is now resolved into a four helix bundle. Structure predictions indicate that only protein IX has a strong propensity for coiled coil formation, thus we have revised our assignment for this exterior density region to a bundle of C-terminal α-helices from protein IX. The early assignment of this density region to protein IIIa is understandable when one considers that the density of four 40-residue α-helices is quite significant (~18 kDa). Before this density region was resolved into α-helices there was no easy way to determine if this represented 18 kDa from a single capsid protein or an assembly of domains from multiple copies of a capsid protein. The recent cryoEM results of Fabry *et al.* (36) on Ad5 and an

adenovirus protein IX deletion mutant actually support this new assignment although it was not clear at the time. They found that when protein IX was deleted, regardless of whether or not the sample was subjected to a single or multiple freeze/thaw cycles, the density previously proposed as protein IIIa at the facet edges was also missing. Our new assignment of the density at the facet edges to a bundle of protein IX C-terminal α-helices is consistent with their published findings.

The second new feature observed at 6 Å resolution is that the large regions of density internal to the penton base are actually clusters of 13 or more long α-helices. Consideration of secondary structure prediction results for all of the capsid protein without atomic resolution structures makes it highly probable that these clusters are N-terminal domains of protein IIIa, which are predicted to have 14 long α-helices. The helical cluster forms extensive interaction surfaces with a peripentonal hexon and with the penton base. Interestingly, a biopanning study using a phage-display peptide library found that a peptide from protein IIIa (aa 157-162 for Ad5) binds to penton base (54). The biopanning result is in agreement with our new assignment for protein IIIa. Also, argon plasma etching experiments point toward a position for protein IIIa between the adenovirus capsid and core (82). Our findings are also in accord with the well established copy number for protein IIIa of 60 copies within the adenovirus virion, or one copy per asymmetric unit (61).

The new assignments for proteins IX and IIIa, which are based on the observation of α-helical density and secondary structure prediction, can be tested biochemically. We have proposed that the C-terminal α-helix of protein IX forms a four-helix bundle and that one α-helix in the bundle interacts with residues in the hexon HVR4 loop. We have

also proposed that the N-terminal two thirds of protein IIIa is likely to fold into a defined domain that interacts with the inner surface of penton base and hexon, and also with an unidentified region of protein VI. Protein IIIa appears to form multiple associations with each of the other proteins in the vertex region of the icosahedral capsid, suggesting a crucial role for protein IIIa as a linchpin for virus assembly and vertex disassembly.

Careful kinetics studies of adenovirus cell entry have indicated that adenovirus particles exist in a metastable state being relatively stable outside of the cell but weakened by the dissociation of structural proteins during the cell entry process so that it can efficiently release its DNA when it reaches the nucleus (47). Wiethoff *et al.* reported that protein IIIa disassociates from the adenovirus capsid at about the same time as penton base and one quarter of the hexons, which are presumably the peripentonal hexons next to penton base (140). Given these previous findings and our new structural information, we propose a more complete model of adenovirus disassembly. When penton base dissociates from the capsid, protein IIIa would also likely be released bringing with it the peripentonal hexons, and multiple copies of protein VI. Protein VI is thought to exist in the virion as a dimer or trimer (18, 35). If multiple copies of protein VI are connected either by direct homotypic association or via interactions with other adenovirus proteins in the core, this would explain the observation by Wiethoff *et al.* that ~80% of protein VI dissociates at the same time as protein IIIa. Protein VI is the adenovirus component that participates in disruption of the endosomal membrane and thus its release from the adenovirus virion is essential for cell entry. We propose that the multiple protein-protein interactions made by protein IIIa in the vertex region facilitate the release of the membrane lytic protein VI.

CHAPTER V

NON-ENVELOPED VIRAL MEMBRANE PENETRATION

Cellular entry processes have been widely studied for many viruses. Understanding these early steps in the viral life cycle is fundamental for the development of vaccines and anti-viral drugs. To fully infect host cells, viruses need access to the cytoplasm or nucleus; therefore, they must penetrate at least one cellular membrane. A common strategy employed by enveloped viruses to access the cellular space is fusion of the viral and a cellular membrane such as the plasma or endosomal membrane. Upon the appropriate trigger such as receptor binding for HIV or low pH for influenza virus, glycoproteins situated in the viral membrane undergo a conformational change and mediate the fusion. On the contrary, non-enveloped viruses lack lipid bilayers, so they cannot breach cellular membranes by a viral-host membrane fusion event. Instead they must employ fundamentally different strategies to breach the cellular barrier. These processes are less well understood than the analogous events for enveloped viruses.

As viruses assemble, they achieve a metastable state primed for cell entry. They must remain stable as they encounter numerous different environments when traveling from host to host. When the virus reaches the appropriate cellular membrane to cross, it must transform itself into a penetration competent state. As with enveloped viruses, non-enveloped viruses need a signal to indicate that the appropriate membrane is near. This signal triggers a conformational change or a cascade of changes that exposes the

penetration protein(s). The membrane is modified in some way that allows the passage of the viral particle or genetic material into the cytoplasm.

Pre- and post-fusion structures of enveloped virus glycoproteins have been well characterized. They are classified according to the presence of an N-terminal fusion peptide in a mainly α-helical glycoprotein (Class I) or a hydrophobic loop within a β-sheet (Class II) (59). Classification of non-enveloped penetration proteins is more challenging because their structure and conformational changes during activation are less understood. Although the structures of many penetration proteins and details of cell entry mechanisms are largely unknown, non-enveloped viruses can be discussed in terms of the types of proteins with known or predicted penetrating capability including fusion-like capsid proteins, myristoylated capsid proteins, capsid proteins with amphipathic helices, and enzymatic capsid proteins. A common theme among non-enveloped viruses is sequestration of the penetration protein until the appropriate time in the viral life cycle. The portion of virus that needs to cross the membrane may influence how penetration proteins interact with host membranes.

Fusion–like penetration proteins

Some penetration proteins show remarkable similarities to the fusion glycoproteins of enveloped viruses. An example is VP4 of rotavirus. Rotavirus is a double-stranded RNA virus and the most prevalent etiological agent of childhood gastroenteritis, which kills approximately 440,000 infants annually (89). A rotavirus vaccine (RotaTeq, Merck & Co., Inc.) is now available against 4 serotypes, but the mechanism of immune protection is not known. The study of the rotavirus cell entry

mechanism including membrane penetration could yield insights into vaccine-based immunity and possibly new vaccines. Rotavirus has a capsid with 3 concentric layers and resembles a spoked-wheel with long VP4 protrusions rooted near the VP6 intermediate capsid layer. There are 120 copies of VP4 in the outer capsid layer. The importance of VP4 is indicated by its participation in cell attachment, penetration, hemagglutination, neutralization, virulence, and protease-enhanced infectivity, hence the interest in understanding its structure (33).

Rotavirus replicates in the enterocytes of the small intestine, an environment abundant in proteases. Trypsin primes rotavirus for infectivity by cleaving VP4 into VP5* and VP8* resulting in the formation of rigid spikes (113, 148). The spikes are bilobed indicating a dimer, but they have pseudohexagonal symmetry at the surface of the outer capsid indicating the presence of a third VP4 molecule in the vicinity. The crystal structure of VP5* missing the anchoring foot (VP5CT) reveals that VP5* can also exist as a trimer (Figure 5.1b) (31). The trimer is stabilized by a triple helix coiled-coil. A β-annulus is formed by the coalescence of strand K from each subunit, and it connects the helix of each subunit to a folded-back globular domain comprised by an 8-stranded antiparallel β sandwich. At the tip of the β sandwich is a hydrophobic apex presumed to interact with the membrane. Interestingly, the antigen domain of VP5*, which lacks the helix and strand K of VP5CT, can pack as a trimer or dimer with essentially the same fold indicated by the low RMSD of 0.85 Å (149). The dimer has both N- and C-termini at the same end of the structure (Figure 5.1a) while VP5CT has its termini at opposing ends indicating a fold back relative to the foot domain, the C-terminus (Figure 5.1b).

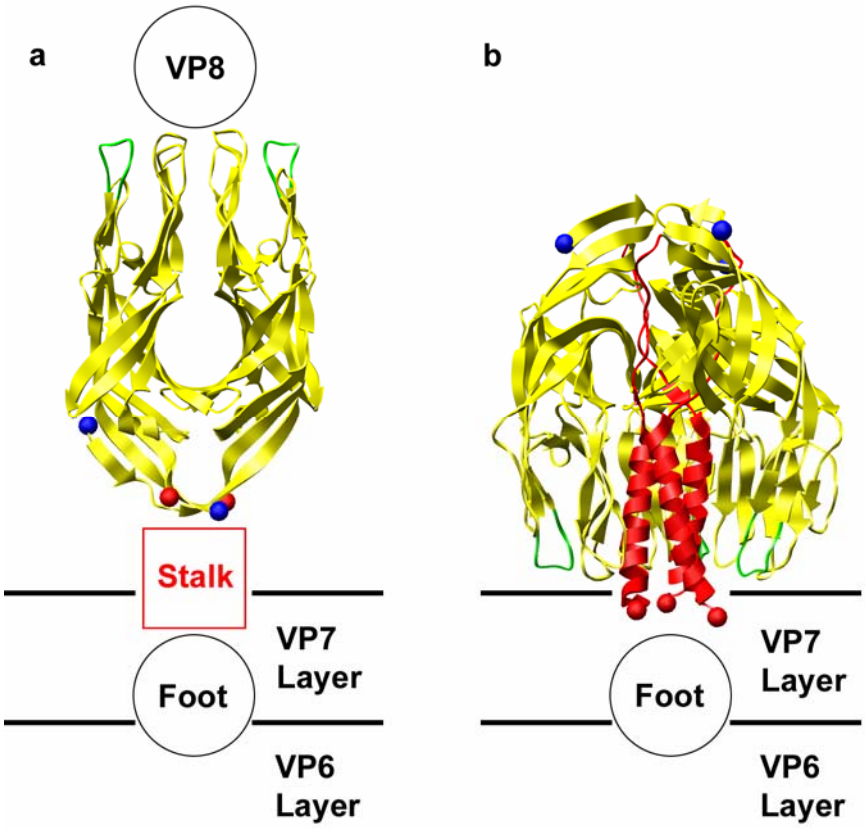


Figure 5.1 Structures of the VP5* protein of rotavirus. **(a)** Crystal structure of the VP5* antigen domain dimer (PDB ID: 2B4H) (149) with the N- and C-termini indicated by blue and red spheres, respectively. The putative membrane interaction loop is in green. The membrane interaction loop is shielded by the companion protein, VP8* (black circle). The remaining stalk (red box) and foot of VP5* (black circle) not crystallized would extend from the C-terminus to the rotavirus capsid. **(a,b)** The two outer concentric capsid layers of VP7 and VP6 are indicated by black lines. **(b)** Crystal structure of the VP5CT trimer (PDB ID: 1SLQ) (31) contains the stalk region of VP5* (aa 480-525) (red). The membrane interaction loop points towards the VP7 capsid layer. The foot domain of VP5* (black circle) would connect to the C-terminus of the stalk. Figure rendered with UCSF Chimera (92).

The dimer is proposed as a "primed" state for cell entry, similar to the dimeric metastable state of class II viral fusion proteins upon budding from cellular membranes (59, 149). A trimer of the VP5* antigen domain was able to form even without the stabilizing coiled coil and β-annulus of VP5CT. A trilobed spike is also observed in the cryoEM structure of rotavirus treated with high pH (91). Thus, the trimer is proposed to be biologically relevant even though the trigger that could induce a dimer to trimer rearrangement with a fold-back of the antigen domain is unknown. This proposed oligomeric rearrangement is strikingly similar to the membrane fusion mechanism used by Semliki Forest virus (alphavirus), dengue virus (flavivirus) and tick-borne encephalitis virus (flavivirus) with class II fusion loops (59). In pre-fusion structures the fusion loop lies at the tip of a β-sheet shielded by a companion protein analogous to VP8* in rotavirus (Figure 5.2a-c) (63, 77, 79, 97, 100, 154). Post-fusion crystal structures of class II fusion proteins indicate that they undergo a dimer to trimer reorganization and a fold-back with a conserved fold upon exposure to low pH (Figure 5.2d-e) which is presumed to bring the two membranes into close proximity (11, 42, 78).

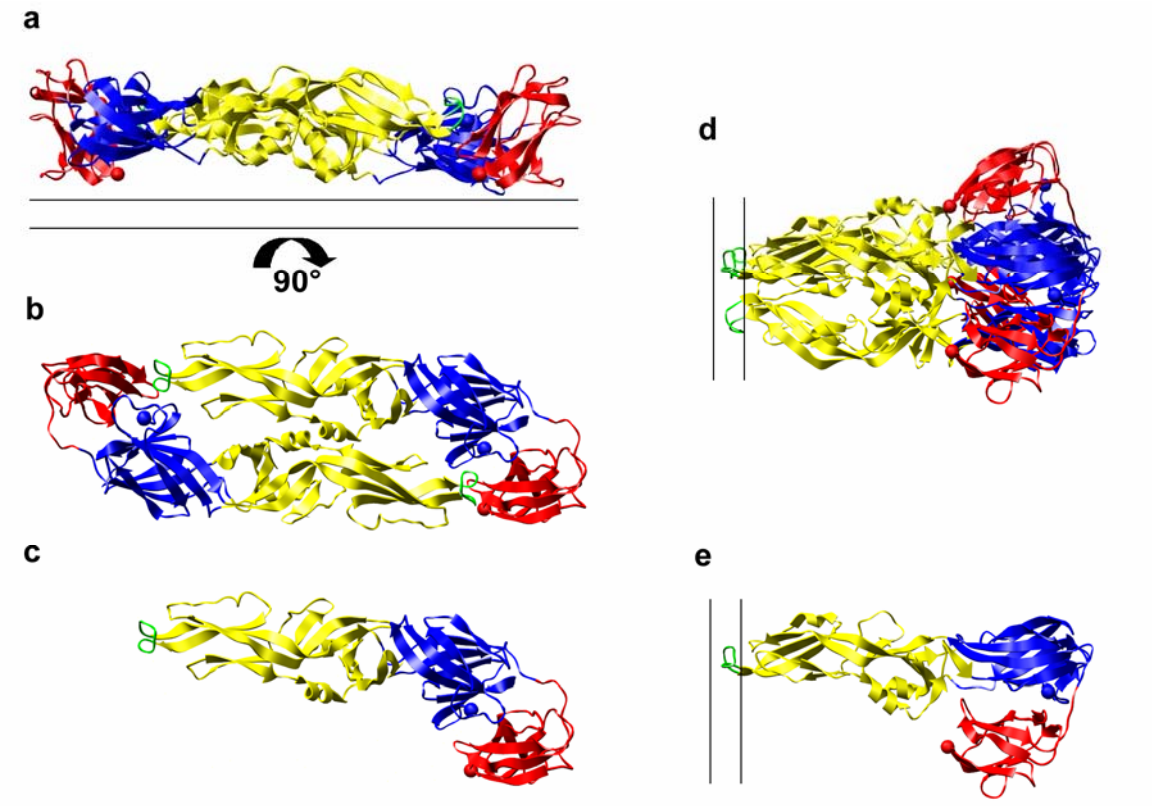


Figure 5.2 Pre- and post-fusion structures of dengue virus E protein. **(a)** Neutral pH (pre-fusion) crystal structure of the soluble ectodomain of the E membrane fusion protein of dengue virus (PDB ID: 1OAN) (77). Residues 395 – 495 containing the putative transmembrane region were not crystallized. Blue, yellow, red, and green represent domains I, II, and III and the fusion loop as defined in (77). The N- and C-termini are indicated by blue and red spheres, respectively. **(a, d, e)** The viral lipid bilayer is represented by 2 black lines. **(b)** View rotated 90° with respect to (a) showing the dimeric oligomeric state of E. **(c)** A monomer of E in the same orientation as (b). **(d)** Low pH (postfusion) crystal structure (PDB ID: 1OK8) (78) of the E as a trimer. **(e)** A monomer of E in the postfusion conformation in the same orientation as (d). Figure rendered with UCSF Chimera (92).

Myristoylation: not merely a coincidence

Non-enveloped viruses in many virus families share the post-translational modification of myristoylation at the N-terminal residue of a capsid protein. Myristoylation involves the substitution of myrisitc acid, a saturated tetradecanoic fatty acid, for the first residue of the polypeptide chain. The myristoyl moiety is believed to insert into the cellular membrane to mediate penetration. Although direct evidence of membrane insertion is lacking in many cases, there is a relationship between myristoylated capsid proteins and cell entry or virus viability. VP4 of poliovirus is one of the best-documented cases.

Poliovirus is a member of the picornavirus family and was one of the most devastating human pathogens of the 20th century. Hence, its structure and mechanisms of cell entry and membrane penetration have been widely studied. The icosahedral capsid contains the typical picornavirus proteins, 60 copies of VP1, VP2, and VP3. Their N-termini create a network with 60 copies of myristoylated VP4 on the inner surface of the capsid near the 5-fold rotational axis (53). In the course of infection the native 160S particle binds to the poliovirus receptor (Pvr), a cell adhesion-like molecule with 3 Ig-like domains (74). This binding induces an irreversible conformational change to the 135S or A-particle. Upon receptor binding, Fricks & Hogle showed that the A particle has reduced affinity for the receptor; the N-terminus of VP1 is externalized; and VP4 is released (39). The 135S particle can then be eluted from membrane fractions, but VP4 remains associated with membranes (29). The externalized N-terminal domain of VP1 has a role in binding to membranes since its proteolytic removal abrogates binding to liposomes (39). A putative amphipathic helix was identified in the N-terminal 20 residues

of VP1, and it was postulated to insert in the membrane to form an ion channel to allow the single-stranded RNA genome passage into the cytoplasm (39).

The initial studies showing a conformational change in VP1 and release of VP4 in response to binding Pvr did not clarify the role of myristoylated VP4 in membrane penetration (39). In the crystal structure of poliovirus, 5 VP4 molecules are situated near the 5-fold rotational symmetry axis, and the myristoyl moiety of one VP4 forms a hydrogen bond with threonine 28 of a nearby VP4 (53). To investigate a potential role of myristoylation in membrane penetration, Moscufo *et al.* constructed VP4 point mutants and discovered that a T28G mutation rendered the virus noninfectious (80). This mutation did not affect the 160S to 135S transition meaning that VP4 could be released from the virion; however, the mutation did cause the particles to be defective in delivering the RNA genome to the cytoplasm (29). Viable mutants with Thr28 substitutions to serine or valine were able to form ion channels in lipid bilayers as wild type 135S particles, but the T28G mutant was defective in ion channel formation. This evidence suggests that VP4 has an active role in cell entry in that VP4 affects the ability of the virus to form ion channels. Thus, VP4 may directly contribute to the architecture of the ion channel.

Polyomavirus is another example in which myristoylation appears to play a role in viral cell entry (69). Polyomavirus is a dsDNA tumor virus, and its capsid is composed 72 pentamers of VP1 (120) with VP3 or myristoylated VP2 associated with the icosahedral 5-fold axis on the inner surface of the virion (20). Following transport to the endoplasmic reticulum (ER), ER protein 29, a possible member of the stress response machinery of the ER, induces a conformational change in VP1 exposing the C-terminus

as observed by mass spectrometry (69). The resulting particle is hydrophobic and can bind to membranes. Virus-like particles containing only VP1 could not associate with liposomes, which implicates another capsid protein in membrane association. The protein is most likely VP2 since mutants devoid of myristoylation have dramatically reduced infectivity, delayed growth, and a defect early in cell entry at or before uncoating, perhaps in membrane association (107). Crystallographic studies suggest a mechanism for VP2 exposure. The VP1 pentamer co-expressed with a fragment of VP2 reveals the C-terminus of VP2 tightly associated with VP1 and the N-terminal portion unresolved and flexible (20). A VP1 rearrangement could allow the extrusion of VP2 giving the flexible N-terminus an avenue to interact with a cellular membrane.

Amphipathic helices in membrane penetration

Human Ads, double-stranded DNA viruses, and poliovirus, a single-stranded RNA virus, both utilize predicted amphipathic helices in membrane penetration. In poliovirus, the predicted amphipathic helix is in the N-terminus of VP1. The N-terminus is important in the early steps of cell entry. If the N-terminus is removed 135S particles cannot attach to liposomes (39), and N-terminal mutations cause delayed RNA release (60). VP4 has been shown to associate with membranes, and VP4 mutations directly impact pore formation and function (29). Whether VP1 directly contributes to pore architecture or anchors the 135S particle to the membrane independently or via VP4, it clearly works in unison with VP4. Even though poliovirus presents a case where a putative amphipathic helix acts in concert with a second penetration protein, mounting

evidence suggests that known or predicted amphipathic helices can be the primary mediators of membrane penetration as is the case for Ad.

The adenovirus protein VI has recently been shown to be critical for membrane penetration (140). Although its atomic structure is not known, protein VI is predicted to have an amphipathic helix in the N-terminal portion of the protein. The role of protein VI in membrane penetration was surprising since the penton base had been suggested to mediate membrane penetration (110, 111). The penton base interacts with α_v integrins to initiate internalization (139). Inhibitory effects of anti-penton base antiserum in the Seth *et al.* study may have interfered with internalization rather than membrane penetration (110). Membrane penetration has been associated with partial capsid disassembly in the endosome (47). Interestingly, a temperature-sensitive mutant of Ad2, *ts*1, cannot escape endosomes and is recycled back to the plasma membrane (46). The P137L mutation in the *ts*1 cysteine protease hinders its encapsidation and activation at the non-permissive temperature (93). As a consequence *ts*1 does not produce mature copies of protein VI and VIII, both internal capsid proteins (137). In an attempt to understand the relationship between disassembly and membrane penetration, Wiethoff *et al*. discovered that *ts*1 is defective in disassembly, and protein VI remains associated with the capsid compared to ~80% release from wt adenovirus virions (140). When protein VI is immunodepleted in a liposome disruption assay, membrane disruption is dramatically reduced. On the other hand, immunodepletion of the penton base has no significant effect on liposome disruption. Recombinant protein VI is also able to mediate lipid disruption and fluorophore release, and the first 54 residues containing the predicted amphipathic helix are indispensable for the lytic activity.

Nemerow and colleagues are carrying out *in vivo* experiments to further substantiate the participation of the putative helix of protein VI in membrane penetration by creating single or double point mutations in the putative amphipathic helix (personal communication, unpublished data). Two of the adenovirus mutant constructs cannot be rescued and therefore may have lethal phenotypes while several other mutants have modestly reduced infectivity as measured by plaque assays. In general, protein VI mutants are properly incorporated into virions, and there are no obvious defects in disassembly. Further experiments are planned to determine whether mutations alter protein VI/lipid interactions.

An animal virus, flock house virus, has a known α-helix in the C-terminus of its capsid protein with multiple helices converging at the 5-fold axis (21). A 44-residue cleavage fragment (γ_1) is autocatalytically derived encompassing the C-terminal α-helix. It was postulated that γ_1 forms a pore in membranes that mediates genome translocation to the cytoplasm. A 21-residue synthetic γ_1 peptide is capable of releasing liposome-bound fluorescent dyes and conducting ions across a membrane (7). Collectively, these results demonstrate that amphipathic helices comprise a third class of penetration proteins.

Enzymatic penetration proteins

A new class of penetration proteins has recently been identified with one known constituent. Parvovirus is a small dsDNA virus composed of 60 copies of only 2 capsid proteins, VP1 and VP2. An active phospholipase A_2 (PLA_2) is encoded in the N-terminal domain of VP1 which is capable of converting phospholipids into fatty acids and other lipophilic substances (152). The N-terminus of VP1 can be exposed *in vitro* by limited

heating (27). Antibodies against the N-terminus of VP1 can block infection when injected into the cytoplasm during parvovirus cell entry suggesting that the VP1 N-terminus is externalized during cell entry (136). A mutation in the PLA_2 domain abrogates lipase activity and renders the virus noninfectious (37). Farr *et al.* were able to increase the infectivity of the mutant by supplying an endosomal rupture agent, polyethyleneimine, or by coinfecting with adenovirus as long as protein VI, the penetration protein of Ad, is released (37). Thus, parvovirus uses a previously unrecognized enzymatic mechanism to pierce a cellular membrane, which is a notable exception to the hydrophobic nature of other classes of penetration proteins.

Penetration proteins are shielded until needed

The particular site of penetration must be coupled to the requirements of the virus life cycle. Exposing the penetration protein too early could be toxic to the host cell before replication begins. If spatial coupling is lost, the penetration protein of a dsDNA virus like polyomavirus could be exposed at the cell surface rather than later in the endoplasmic reticulum, which would probably interfere with genome delivery to the nucleus for replication. An emerging theme in membrane penetration by non-enveloped viruses is the sequestration of the penetration protein until the preferred cellular membrane is near.

The membrane penetrating ability of reovirus is thought to originate from myristoylated μ1 (85), a major outer capsid protein present in 600 copies. Even though μ1 is the major protein on the outside of the capsid (Figure 5.3a-c), Reovirus has devised a clever strategy to shield the myristoyl moiety until the release trigger is received. In the

atomic resolution structure of μ1 complexed with its protector protein, σ3, a hydrophobic pocket was found on the side of the μ1 trimer (1 pocket/monomer) which contained the crystallization detergent (64). Liemann *et al.* proposed that the myristoyl group occupies this pocket in the intact virion. The first eight residues were disordered, so additional information is needed to address this hypothesis. In a recent cryoEM analysis of reovirus at 7.0 Å resolution, Zhang *et al*. observed a tube of density extending from Thr10 of the docked crystal structure to the hydrophobic binding pocket observed by Liemann *et al*. consistent with the idea of the N-terminus reaching to this pocket and shielding the hydrophobic myristoyl group (Figure 5.3d-e) (153).

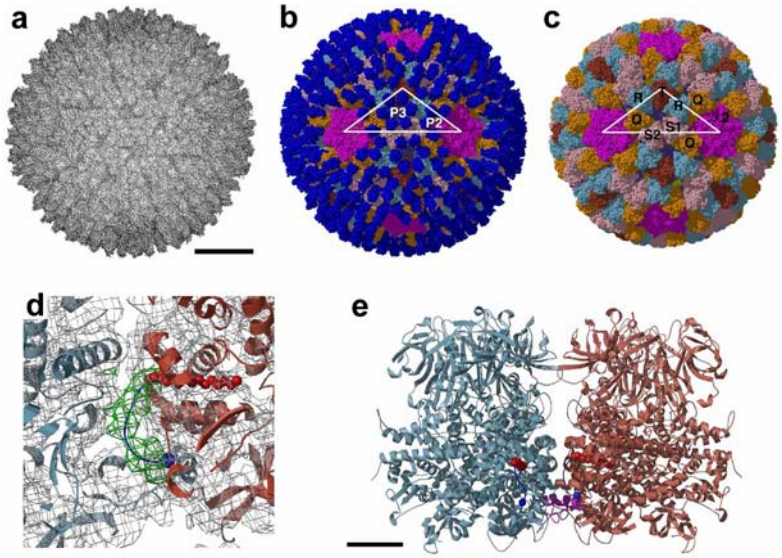


Figure 5.3 Reovirus cryoEM density at 7Å resolution. **(a)** CryoEM structure of the reovirus virion viewed along the 2-fold rotational axis. **(a,b,c)** The scale bar represents 200 Å. **(b and c)** Pseudoatomic models of the virion (b) and ISVP **(c)** generated by docking the μ1/σ3 heterohexamer (PDB ID: 1JMU) (64) and λ2 (PDB ID: 1EJ6) (95) **(b)** or λ2 and the μ1 hexamer only (c) into the cryoEM density. λ2 pentamers are in magenta, σ3 in blue, and μ1 trimers in cyan, pink, orange, and brown. The asymmetric unit is outlined by a white triangle and P2 and P3 channels (b) and capsid proteins (c) are labeled. **(d)** Close-up view of cryoEM density (mesh) and 2 docked μ1 trimers. A myristoyl group (red) is modeled in the hydrophobic pocket identified in (64). Extra cryoEM density (green mesh) has been assigned to aa 2 – 9 and a preliminary backbone trace for these residues is shown in blue. Thr10 is marked by a blue sphere. **(e)** Demagnified view in the same orientation as (d) showing a modeled myristoyl group in each trimer. Also a tentative trace for residues 72 – 96 is shown (magenta). Scale bar, 25 Å. Figure reprinted from (153) and modified with permission from Elsevier.

As with adenovirus, reovirus undergoes major structural changes during endocytosis. The outer capsid protein, σ3, is lost after cleavage by lysosomal acidic proteases, which generates infectious subvirion particles (ISVPs). ISVPs, not virions or cores, have membrane penetrating activity (129). The formation of ISVPs is followed by dissociation of the fiber protein (σ1), and a conformational change in μ1 renders ISVPs hydrophobic. The hydrophobic form of ISVP has been proposed as an additional cell entry intermediate called ISVP* (16, 17, 86). During cell entry, μ1 is autocatalytically cleaved between Asn42 and Pro43, which produces 4 kDa (μ1N) and 72 kDa (μ1C) fragments (37, 84). Cleavage of Asn/Pro43 appears to occur after the conformational change in μ1 because a μ1N/μ1C cleavage mutant, $\mu1^{N42A}$, contained the hydrophobic form of μ1 (86). Although the ISVP → ISVP* transition has not been characterized structurally, it could involve extracting the myristoyl group from the binding pocket and setting up for the Asn42/Pro43 cleavage. The μ1N fragment is normally released following cleavage. The $\mu1^{N42A}$ mutant retains the μ1N fragment and is not able to permeabilize membranes *in vitro* or *in vivo* indicating that the generation of μ1N and its release are important in cell entry (84, 86).

The 6 Å resolution cryoEM reconstruction of Ad5 pseudotyped with the short fiber from Ad35 (Ad35F) suggests that the adenovirus penetration protein, protein VI, may be shielded in a fashion similar to that of reovirus μ1. Density in the cavity of the hexon has been assigned to protein VI based upon proximity to hexon residue 776, which has been implicated in protein VI binding (141). There are three symmetry-related density rods that are within 10 Å of the three copies of hexon residue 776 (Figure 4.6). Some of the density rods extend and fold over the base of the hexon (Figures 4.6c and

5.4a). Both the density rod within the hexon cavity and a portion of the region folded over the base of the hexon are long enough to accommodate ~3 helical turns (Figure 5.4b,c). These could be the predicted α-helices of the N-terminal region of protein VI. Both predicted α-helices overlap the N-terminal hexon binding domain (hbd1) (Figure 5.4d). Hbd1 includes residues 48 – 74 which is about 3 kDa in molecular mass (72). At a moderate isosurface level, the volume of the protein VI rod and the protein VI density folded over the base of hexon corresponds to ~2.5 kDa in molecular mass, which is consistent with the size of hbd1. Therefore, the density in the cavity of the hexon is tentatively assigned to the N-terminus of protein VI including the membrane lytic region. In this model, the nuclear localization signal, NLS1 (aa 131 – 135) (Figure 5.4), does not fall in the region of protein VI bound to the hexon cavity. Since one of the functions of protein VI in the adenovirus life cycle is shuttling hexon to the nucleus for viral assembly (141), NLS1 and NLS2 must be accessible while protein VI is associated with hexon. This model suggests that the N-terminal region of protein VI is shielded in the hexon cavity until its membrane lytic activity is needed in the life cycle.

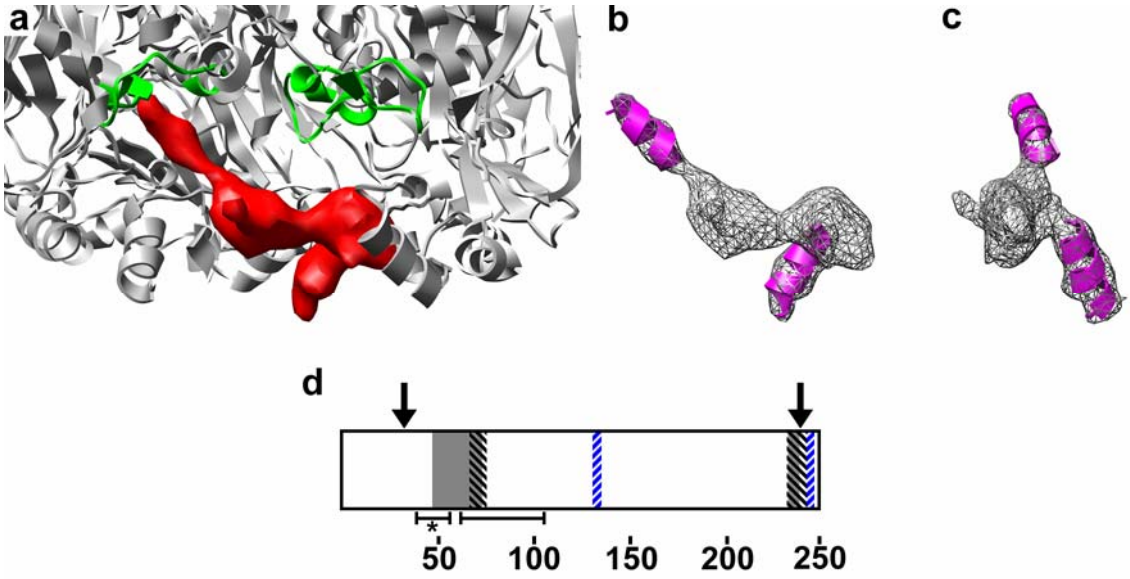


Figure 5.4 Density within the cavity of the hexon trimer assigned as the N-terminal domain of protein VI. **(a)** Region of Fig 4.6c magnified to show only the protein VI density rod and density folded over the base of the hexon (red). The hexon coordinates are shown in gray ribbons and loop implicated in protein VI binding (aa 770 – 787) is shown in green. **(b)** Protein VI density (gray mesh) fitted with 9- and 14- residue helices (magenta). **(c)** View as in **(b)** with a 90 degree rotation. **(d)** Schematic bar diagram of protein VI. Hexon binding domains 1 and 2 are shown in gray, and nuclear localization and export signals are marked by blue and black diagonal lines, respectively. Predicted helices are shown by brackets, with the amphipathic helix marked by an asterisk. (See Table 4.4). Protease cleavage sites at residues 33 and 239 are marked by arrows.

Penetration protein interactions with membranes

Although crystal structures are available for poliovirus (53), μ1 of reovirus (64), and fragments of myristoylated VP2 from polyomavirus (20) among others, high resolution structures for many penetration proteins have not been determined. Solving structures of penetration proteins solubilized under penetrating conditions or in the context of an intact membrane presents a daunting challenge. Thus, the precise events in membrane penetration by non-enveloped viral proteins have not been adequately addressed. With the limited knowledge to date it is surmised that penetration proteins form discrete pores of limited size in the case of rhinovirus (10), large membrane disturbances in the case of adenovirus (10), or perhaps induce chaperone-assisted transport in the case of polyomavirus (69). Since non-enveloped viruses can be single- or double-stranded DNA or RNA viruses, what needs to cross the cellular membrane and the membrane to be traversed differs. How the penetration protein interacts with the cellular membrane is most likely dictated by the constraints of the virus life cycle.

Poliovirus has a single, plus-stranded RNA genome meaning that the genome is ready for translation in the cytoplasm and must produce the virally encoded RNA-synthesizing machinery to make minus strands for replication. Even though the exact site of penetration has not been elucidated, only the RNA needs to cross the membrane for a viable infection to ensue. If only the RNA needs passage, then a large membrane disturbance would not be necessary. Rhinovirus is also a member of the picornavirus family. In endosome leakage experiments rhinovirus allowed the release of only small dextran molecules consistent with the idea of a small pore that would facilitate transport of RNA, which is approximately 10 Å in diameter (10). Experiments confirm that

poliovirus cell entry intermediates form pores in planar lipid bilayers of about 20 Å in diameter (128) and that VP4 is most likely involved in pore formation (29).

Reovirus has markedly different needs for replication. Reovirus relies on the transcriptionally active core to produce mRNA transcripts from the dsRNA genome followed by mRNA liberation from the 5-fold vertex. It is not clear when the transition from ISVP to transcriptionally active core particle is made, but at least the core needs access to the cytoplasm where transcription begins. The breach in the ~100 nm endosomal vesicle must be at least ~70 nm in diameter to allow cores to pass unobstructed. One study reported the formation of multisized channels in planar lipid bilayers; however, the size distribution of such channels was not clear including the size that could be formed *in vivo* (129). Recently, it was discovered that μ1N must be released to permeabilize membranes (86), and if 600 copies are released simultaneously the membrane could be carpeted and ruptured for release of the core particle. Although channel formation has been reported *in vitro*, a large membrane disturbance should not be ruled out as the mechanism by which reovirus exits the late endosome.

Since adenovirus is a double-stranded DNA virus, its genome must replicate in the nucleus. Adenovirus particles have been detected in the cytoplasm and even docked at the nuclear pore complex (45). The ~93 nm capsid must also exit endosomal vesicles of ~100 nm in diameter. Thus, a large membrane disturbance seems more likely than channel formation. Data supporting this notion comes from experiments analyzing the release of different size dextran molecules from endocytic vacuoles (10). Following the dissociation of the penton base and protein IIIa in the acidified endosome, nearly 300 copies of protein VI are deployed (140). The concentration of protein VI in the endosome

is approximately 2 mg/ml (G.Nemerow, personal communication). It seems plausible that this massive release of protein VI could carpet and rupture the endosomal membrane allowing the partially disassembled virion access to the cytoplasm and microtubules for transport to the nucleus.

Conclusions

Non-enveloped viruses have devised clever strategies to penetrate host cell membranes to gain access to the cytoplasm. Both structural and functional analyses have indicated that membrane penetration by non-enveloped viruses can involve a fusion-like mechanism as is the case for rotavirus, myristoylated capsid proteins for viruses like polyomavirus and reovirus, or a capsid protein with an amphipathic helix like in Ad. The discovery of enzymatic membrane lysis by an NEV has generated a new class of penetration proteins that rely on characteristics significantly different than penetration proteins with hydrophobic domains in other classes. A common feature of all classes is the protection of the lytic protein or domain until membrane penetration is needed. The interactions between non-enveloped viral proteins and membranes appear to be related to the cytoplasmic passage of viral genetic material, a transcription complex, or a viral capsid.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Adenovirus has been at the heart of gene therapy development as evidenced by the numerous clinical trials in progress using adenovirus-based vectors as therapeutic agents (52). Adenovirus-based gene therapy is promising, but there are obstacles to overcome in the process of bringing therapies to the bedside. Most prominent is the activation of the host immune response when delivering first generation vectors intravenously. Major efforts are underway to retarget adenovirus vectors away from CAR-expressing tissues and towards other desired targets. Ad35, a subgroup B virus, shows great promise as a vector due to its ability to infect non-CAR expressing cells. In preliminary mouse *in vivo* studies the Ad35F vector (117), which is composed of the Ad5 capsid pseudotyped with the Ad35 fiber, primarily infects tumor tissue and evades immune system activation typically seen with Ad5 vectors (83). The absence of a high resolution structure of adenovirus has hampered further vector development, and this has been a motivating factor in the work presented in this dissertation. This chapter summarizes the significance of solving subnanometer resolution structures of adenovirus, the impact of visualizing α-helices on protein assignments, implications of this work for viral assembly and capsid disassembly, and possible future adenovirus structural studies.

Significance of determining subnanometer resolution structures of Ad

The major challenge of imaging biological complexes with an electron microscope is that proteins suffer radiation damage induced by the electron beam. To compensate for the low signal-to-noise ratio in two-dimensional particle images, large data sets on the order of $10^4 – 10^5$ particle images are averaged. The literature indicates that α-helices can be resolved when the structure reaches ~10 Å resolution (56), so our first goal was to reach this resolution level for the adenovirus reconstruction. There are a growing number of cryoEM structures in the subnanometer resolution range (56), but only one subnanometer resolution structure has been determined for a complex as large or larger than adenovirus: herpesvirus at 8.5 Å resolution (156). Herpesvirus has an icosahedral capsid of 1250 Å in diameter surrounded by a proteinaceous tegument layer and a glycoprotein envelope. At 8.5 Å resolution, it was observed that the distribution α-helices in the VP5 and VP23 capsid proteins did not match any in known viral structures, indicating that VP5 and VP23 may have unique folds.

It has been predicted that 100,000 views of the asymmetric unit are necessary to reach 10 Å resolution and approximately 1 million to reach atomic (3 Å) resolution (43). The number of particles can be reduced if the complex of interest has symmetry. For example, adenovirus is an icosahedral virus with 60-fold symmetry, and thus the predicted number of images necessary to reach 10 Å resolution is 1,667 (100,000/60 = 1,667) and to reach atomic resolution is 16,667 (1,000,000/60 = 16,667). We were able to achieve 9 Å resolution for Ad35F reconstruction in Aim I with only 960 particle images, significantly less than the predicted number. In Aim II we were able to refine the Ad35F structure to 6 Å resolution with just over 2,600 particle images, or 156,000 views of the

asymmetric unit (2,600 x 60 = 156,000). For comparison the cryoEM structure of GroEL at 6 Å resolution (67) was produced with 39,000 particle images. Considering the 14-fold symmetry of GroEL this corresponds to 546,000 views of the asymmetric unit (39,000 x 14 = 546,000), which is over three times the number of views of the asymmetric unit than we used in the 6 Å resolution structure of Ad35F. These results imply that processing with the highest quality data may make the daunting challenge of achieving atomic resolution by cryoEM imaging more manageable.

Impact of visualizing α-helices on protein assignments

We have shown that cryoEM can provide sufficient resolution of a viral structure to resolve α-helices. In the Aim I structure of Ad35F at 9 Å resolution, we were able to visualize a long ~40 residue α-helix, which we assigned to the C-terminus of protein IX. The shorter α-helices of the other capsid proteins without atomic resolution structures were not observed, so in Aim II we refined the Ad35F reconstruction to 6 Å resolution. In the Aim II structure we were able to resolve α-helices of 10 residues or longer. After fitting the crystal structures of hexon and penton base into the cryoEM reconstruction, we confirmed that each α-helix of at least 10 residues (3 or more turns) in the crystal structures was resolved as a rod in the cryoEM density map.

Achieving the milestone of resolving α-helices allowed us to re-evaluate the density assignments made in the first adenovirus difference map at 25 Å resolution (123). This was the first demonstration that electron density maps are accurate enough to allow crystal structures to be docked into their corresponding locations in the EM envelope. In 1993 only the hexon crystal structure was available for this study, and it was docked into

240 positions in the adenovirus capsid and subtracted to reveal density for the other capsid proteins. The difficulty in interpreting a 25 Å resolution difference map of a complex viral capsid is discerning whether density regions represent individual proteins or multimers of domains from multiple proteins. For this reason the 60 tetrameric coiled coils of protein IX were incorrectly assigned to 60 copies of protein IIIa. It is now clear that all of the density in addition to hexon, penton base, and fiber on the outer surface of the virion is protein IX rather than a combination of protein IX and protein IIIa. This implies that all the other capsid proteins are internal. Visualizing α-helices also enabled us to make another major density reassignment for inner capsid density. The density formerly assigned to multimers of protein VI at 25 Å resolution is now resolved into a helical cluster closely associated with vertex proteins on the inner capsid surface. Comparison of the number and length of observed α-helices with secondary structure predictions for proteins IIIa, VI, and VIII made clear that the cluster is protein IIIa. The proximity of protein IIIa to protein VI density nestled in the adjacent hexon cavity has important implications for capsid disassembly.

Implications for viral assembly

Protein IX has been known as the capsid 'cement' since protein IX mutant virions are more heat labile and do not form characteristic groups of nine (GON) hexons when dissociated. Protein IX and hexon are a few of earliest components incorporated into progeny virions together with the penton base and scaffolding proteins, which are subsequently released and are not found in mature virions (44). The 9 Å resolution structure of Ad35F revealed that the four trimeric regions of density observed in each

icosahedral facet interacting with the GON hexons are most likely trimers of only the N-terminal domain of protein IX. Our findings suggest that the strict requirement for the conserved N-terminal domain of protein IX for incorporation into virions during assembly (98) is a result of the N-terminus of protein IX interacting with the GON hexons.

The refinement of the protein IIIa position helps clarify the known role of protein IIIa in viral assembly. Temperature-sensitive mutants with defects in DNA encapsidation have mutations in protein IIIa indicating that protein IIIa has a role in packaging the viral genome (25). With the prior assignment of protein IIIa on the outer capsid surface and a small portion on the inner capsid surface, it was difficult to envision how protein IIIa could facilitate DNA packaging (123). Biochemical studies indicate that protein IIIa crosslinks to protein VII (34), a core protein known to interact with the viral DNA. The assignment of the tetrameric coiled coil on the outer capsid surface to protein IX and the reassignment of protein IIIa to the cluster of helices associated with vertex proteins on the inner capsid surface placed protein IIIa entirely in the interior of the particle. The C-terminal one-third of protein IIIa, core proteins, and the viral DNA were not observed in the reconstruction suggesting that protein IIIa interacts with protein VII via its C-terminal region. This interaction would effectively link the capsid with the core proteins and viral DNA, which could facilitate packaging of the genome during viral assembly.

Implications for capsid disassembly during cell entry

The reassignment of protein IIIa density has also allowed us to form new hypotheses on capsid disassembly. Protein VI participates in disruption of the endosomal membrane, and its release from the adenovirus virion is essential for cell entry. Protein VI is sequestered in the interior of the virus probably because it is toxic to cells, and it would be advantageous to keep it shielded until its lytic activity is needed during cell entry. The cryoEM structure suggests that the membrane lytic predicted amphipathic helix is shielded in the hexon cavity. In adenovirus, it is known that low pH triggers the liberation of the penton base, about a quarter of the hexons, and the majority of protein IIIa and VI. Prior to this work it was not evident how a significant percentage of protein VI could be efficiently released for endosomal membrane penetration. The newly assigned protein IIIa density is closely associated with the penton base, 25% of protein VI, and 25% of hexons. The cryoEM structure suggests that the nearly simultaneous release of these proteins during cell entry may be a result of the extensive protein/protein interfaces observed between protein IIIa and both hexon and penton base. Biochemical data suggest that protein VI forms trimers or dimers (18). If protein VI forms a network within the virion, then rapid vertex dissociation facilitated by protein IIIa would provide an efficient avenue for the release of a significant percentage of protein VI from the virion. In addition to a role for protein IIIa in DNA encapsidation during virion assembly, we propose that protein IIIa serves as the linchpin for vertex dissociation in cell endosomes to facilitate the release of the membrane lytic protein VI.

Future Adenovirus Structural Studies

The Ad35F cryoEM structures at 6 and 9 Å resolution have allowed refinements in the assignment of capsid density; however, a higher resolution structure could further enhance our understanding of the capsid. At 6 Å resolution α-helices were visualized in the adenovirus capsid, but the resolution was not great enough to identify large side chains. Thus, a unique fold could not be proposed for protein IIIa. Achieving 4.5 or 3 Å resolution cryoEM single particle reconstruction would be a landmark achievement not only for adenovirus, but also for the cryoEM community. In a cryoEM helical reconstruction at 4 Å resolution of flagellin, which forms the bacterial flagellar filament, the backbone was traced for 100 terminal residues that were missing from the F41 fragment (150). In the core of the filament, the path of the α-helical backbone and some large side chains were visible. By refining the 6 Å resolution Ad35F structure to 4.5 Å resolution it may be possible to resolve large, bulky side chains and begin to trace the fold for capsid proteins IIIa, VI, VIII, and IX. Once large sidechains are resolved in hexon and penton base, then bulky residues could be visualized in α-helices thereby giving directionality to the peptide backbone. Visualization of large sidechains in the protein IIIa density would be particularly useful in *de novo* fold modeling studies underway in collaboration with Jens Meiler's laboratory. At 6 Å resolution, modeling techniques could not differentiate between 2000 possible models for the structure of protein IIIa. Sidechain information from a 4.5 Å resolution structure could help reduce the candidate models in order to converge upon the unique fold for protein IIIa. Likewise, visualization of bulky sidechains in the protein VI density, could help confirm that the N-terminal membrane lytic region of protein VI is shielded within the hexon cavity.

The resolution of the Aim II structure of Ad35F leveled off at 6.9 Å using a conservative estimate (0.5 Fourier Shell Correlation criterion), and suggests how the structure could be refined to higher resolution. Approximately 65% of the images reconstructed originated from dataset 1, which had a final pixel size of 2.4 Å. By the Nyquist Theorem it would be theoretically possible to achieve 4.8 Å resolution (two times the pixel size of 2.4 Å), but a resolution of three times the pixel size is more likely, 7.2 Å resolution. The remaining 35% of the particle images had a pixel size of 1.5 Å (dataset 3), and by including them in the reconstruction the resulting resolution was better than three times the pixel size of dataset 1. This implies that collecting and processing more images like dataset 3 would make refinement to 4.5 Å resolution possible with the Frealign image processing software package. Furthermore, processing images similar to dataset 3 and binned by three to a pixel size of 1.1 Å could make atomic (~ 3 Å) resolution achievable for adenovirus by cryoEM.

Concluding remarks

These structural studies using cryoEM single particle reconstruction techniques have greatly enhanced our knowledge of the adenovirus capsid architecture. In Aim I we determined a 9 Å resolution structure of adenovirus and this enabled us to visualize a 40-residue long α-helix in the capsid. In Aim II we determined a 6 Å resolution structure of adenovirus and this enabled us to resolve all α-helices of 10 or more residues in the cryoEM density. Resolving α-helices has been pivotal in understanding where proteins are located in the capsid. The Aim II structure also suggested that protein IIIa may have

an additional role in the adenovirus life cycle. The new knowledge of the locations of capsid proteins will likely pave the way for future vector retargeting approaches.

REFERENCES

1. **Adiga, U., W. T. Baxter, R. J. Hall, B. Rockel, B. K. Rath, J. Frank, and R. Glaeser.** 2005. Particle picking by segmentation: a comparative study with SPIDER-based manual particle picking. J. Struct. Biol. **152:**211-20.

2. **Adrian, M., J. Dubochet, J. Lepault, and A. W. McDowall.** 1984. Cryo-electron microscopy of viruses. Nature **308:**32-6.

3. **Akalu, A., H. Liebermann, U. Bauer, H. Granzow, and W. Seidel.** 1999. The subgenus-specific C-terminal region of protein IX is located on the surface of the adenovirus capsid. J Virol **73:**6182-7.

4. **Berger, B., D. B. Wilson, E. Wolf, T. Tonchev, M. Milla, and P. S. Kim.** 1995. Predicting coiled coils by use of pairwise residue correlations. Proc Natl Acad Sci U S A **92:**8259-63.

5. **Bewley, M. C., K. Springer, Y. B. Zhang, P. Freimuth, and J. M. Flanagan.** 1999. Structural analysis of the mechanism of adenovirus binding to its human cellular receptor, CAR. Science **286:**1579-83.

6. **Bischoff, J. R., D. H. Kirn, A. Williams, C. Heise, S. Horn, M. Muna, L. Ng, J. A. Nye, A. Sampson-Johannes, A. Fattaey, and F. McCormick.** 1996. An adenovirus mutant that replicates selectively in p53-deficient human tumor cells. Science **274:**373-6.

7. **Bong, D. T., C. Steinem, A. Janshoff, J. E. Johnson, and M. Reza Ghadiri.** 1999. A highly membrane-active peptide in Flock House virus: implications for the mechanism of nodavirus infection. Chem Biol **6:**473-81.

8. **Borgnia, M. J., D. Shi, P. Zhang, and J. L. Milne.** 2004. Visualization of alpha-helical features in a density map constructed using 9 molecular images of the 1.8 MDa icosahedral core of pyruvate dehydrogenase. J Struct Biol **147:**136-45.

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

10. **Brabec, M., D. Schober, E. Wagner, N. Bayer, R. F. Murphy, D. Blaas, and R. Fuchs.** 2005. Opening of size-selective pores in endosomes during human rhinovirus serotype 2 in vivo uncoating monitored by single-organelle flow analysis. J Virol **79:**1008-16.

11. **Bressanelli, S., K. Stiasny, S. L. Allison, E. A. Stura, S. Duquerroy, J. Lescar, F. X. Heinz, and F. A. Rey.** 2004. Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. Embo J **23:**728-38.

12. **Bryson, K., L. J. McGuffin, R. L. Marsden, J. J. Ward, J. S. Sodhi, and D. T. Jones.** 2005. Protein structure prediction servers at University College London. Nucleic Acids Res. **33:**W36-8.

13. **Burmeister, W. P., D. Guilligay, S. Cusack, G. Wadell, and N. Arnberg.** 2004. Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. J. Virol. **78:**7727-36.

14. **Campos, S. K., M. B. Parrott, and M. A. Barry.** 2004. Avidin-based targeting and purification of a protein IX-modified, metabolically biotinylated adenoviral vector. Mol Ther **9:**942-54.

15. **Chacón, P., and W. Wriggers.** 2002. Multi-resolution contour-based fitting of macromolecular structures. J. Mol. Biol. **317:**375-84.

16. **Chandran, K., D. L. Farsetta, and M. L. Nibert.** 2002. Strategy for nonenveloped virus entry: a hydrophobic conformer of the reovirus membrane penetration protein micro 1 mediates membrane disruption. J Virol **76:**9920-33.

17. **Chandran, K., J. S. Parker, M. Ehrlich, T. Kirchhausen, and M. L. Nibert.** 2003. The delta region of outer-capsid protein micro 1 undergoes conformational change and release from reovirus particles during cell entry. J Virol **77:**13361-75.

18. **Chatterjee, P. K., M. E. Vayda, and S. J. Flint.** 1985. Interactions among the three adenovirus core proteins. J. Virol. **55:**379-86.

19. **Chee, A. V., P. Lopez, P. P. Pandolfi, and B. Roizman.** 2003. Promyelocytic leukemia protein mediates interferon-based anti-herpes simplex virus 1 effects. J Virol **77:**7101-5.

20. **Chen, X. S., T. Stehle, and S. C. Harrison.** 1998. Interaction of polyomavirus internal protein VP2 with the major capsid protein VP1 and implications for participation of VP2 in viral entry. Embo J **17:**3233-40.

21. **Cheng, R. H., V. S. Reddy, N. H. Olson, A. J. Fisher, T. S. Baker, and J. E. Johnson.** 1994. Functional implications of quasi-equivalence in a T = 3 icosahedral animal virus established by cryo-electron microscopy and X-ray crystallography. Structure **2:**271-82.

22. **Chiu, C. Y., P. Mathias, G. R. Nemerow, and P. L. Stewart.** 1999. Structure of adenovirus complexed with its internalization receptor, αvβ5 integrin. J. Virol. **73:**6759-68.

23. **Chiu, C. Y., E. Wu, S. L. Brown, D. J. Von Seggern, G. R. Nemerow, and P. L. Stewart.** 2001. Structural analysis of a fiber-pseudotyped adenovirus with ocular tropism suggests differential modes of cell receptor interactions. J. Virol. **75:**5375-80.

24. **Chroboczek, J., R. W. Ruigrok, and S. Cusack.** 1995. Adenovirus fiber. Curr. Top. Microbiol. Immunol. **199 (Pt 1):**163-200.

25. **Chroboczek, J., F. Viard, and J. C. D'Halluin.** 1986. Human adenovirus 2 temperature-sensitive mutant 112 contains three mutations in the protein IIIa gene. Gene **49:**157-160.

26. **Conway, J. F., N. Cheng, A. Zlotnick, P. T. Wingfield, S. J. Stahl, and A. C. Steven.** 1997. Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. Nature **386:**91-4.

27. **Cotmore, S. F., M. D'Abramo A, Jr., C. M. Ticknor, and P. Tattersall.** 1999. Controlled conformational transitions in the MVM virion expose the VP1 N-terminus and viral genome without particle disassembly. Virology **254:**169-81.

28. **Crawford-Miksza, L., and D. P. Schnurr.** 1996. Analysis of 15 adenovirus hexon proteins reveals the location and structure of seven hypervariable regions containing serotype-specific residues. J. Virol. **70:**1836-44.

29. **Danthi, P., M. Tosteson, Q. H. Li, and M. Chow.** 2003. Genome delivery and ion channel properties are altered in VP4 mutants of poliovirus. J Virol **77:**5266-74.

30. **Deane, C. M., and T. L. Blundell.** 2001. CODA: a combined algorithm for predicting the structurally variable regions of protein models. Protein Sci **10:**599-612.

31. **Dormitzer, P. R., E. B. Nason, B. V. Prasad, and S. C. Harrison.** 2004. Structural rearrangements in the membrane penetration protein of a non-enveloped virus. Nature **430:**1053-8.

32. **Durmort, C., C. Stehlin, G. Schoehn, A. Mitraki, E. Drouet, S. Cusack, and W. P. Burmeister.** 2001. Structure of the fiber head of Ad3, a non-CAR-binding serotype of adenovirus. Virology **285:**302-12.

33. **Estes, M. K.** 1996. Rotaviruses and Their Replication, p. 1625 - 1655. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Fields Virology, Third ed, vol. 2. Lippincott - Raven Publishers, Philadelphia.

34. **Everitt, E., L. Lutter, and L. Philipson.** 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.

35. **Everitt, E., and L. Philipson.** 1974. Structural proteins of adenoviruses. XI. Purification of three low molecular weight virion proteins of adenovirus type 2 and their synthesis during productive infection. Virology **62:**253-69.

36. **Fabry, C. M., M. Rosa-Calatrava, J. F. Conway, C. Zubieta, S. Cusack, R. W. Ruigrok, and G. Schoehn.** 2005. A quasi-atomic model of human adenovirus type 5 capsid. EMBO J. **24:**1645-54.

37. **Farr, G. A., L. G. Zhang, and P. Tattersall.** 2005. Parvoviral virions deploy a capsid-tethered lipolytic enzyme to breach the endosomal membrane during cell entry. Proc Natl Acad Sci U S A **102:**17148-53.

38. **Fishelson, Z., N. Donin, S. Zell, S. Schultz, and M. Kirschfink.** 2003. Obstacles to cancer immunotherapy: expression of membrane complement regulatory proteins (mCRPs) in tumors. Mol Immunol **40:**109-23.

39. **Fricks, C. E., and J. M. Hogle.** 1990. Cell-induced conformational change in poliovirus: externalization of the amino terminus of VP1 is responsible for liposome binding. J Virol **64:**1934-45.

40. **Furcinitti, P. S., J. van Oostrum, and R. M. Burnett.** 1989. Adenovirus polypeptide IX revealed as capsid cement by difference images from electron microscopy and crystallography. EMBO J. **8:**3563-70.

41. **Gaggar, A., D. M. Shayakhmetov, and A. Lieber.** 2003. CD46 is a cellular receptor for group B adenoviruses. Nat. Med. **9:**1408-12.

42. **Gibbons, D. L., M. C. Vaney, A. Roussel, A. Vigouroux, B. Reilly, J. Lepault, M. Kielian, and F. A. Rey.** 2004. Conformational change and protein-protein interactions of the fusion protein of Semliki Forest virus. Nature **427:**320-5.

43. **Glaeser, R. M.** 1999. Review: electron crystallography: present excitement, a nod to the past, anticipating the future. J Struct Biol **128:**3-14.

44. **Greber, U. F.** 1998. Virus assembly and disassembly: the adenovirus cysteine protease as a trigger factor. Rev Med Virol **8:**213-222.

45. **Greber, U. F., M. Suomalainen, R. P. Stidwill, K. Boucke, M. W. Ebersold, and A. Helenius.** 1997. The role of the nuclear pore complex in adenovirus DNA entry. Embo J **16:**5998-6007.

46. **Greber, U. F., P. Webster, J. Weber, and A. Helenius.** 1996. The role of the adenovirus protease on virus entry into cells. Embo J **15:**1766-77.

47. **Greber, U. F., M. Willetts, P. Webster, and A. Helenius.** 1993. Stepwise dismantling of adenovirus 2 during entry into cells. Cell **75:**477-86.

48. **Grigorieff, N.** 2006. FREALIGN: High-resolution refinement of single particle structures. J Struct Biol. **in press**

49. **Grigorieff, N.** 1998. Three-dimensional structure of bovine NADH:ubiquinone oxidoreductase (complex I) at 22 A in ice. J. Mol. Biol. **277:**1033-46.

50. **Guex, N., and M. C. Peitsch.** 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis **18:**2714-23.

51. **Habib, N., H. Salama, A. Abd El Latif Abu Median, I. Isac Anis, R. A. Abd Al Aziz, C. Sarraf, R. Mitry, R. Havlik, P. Seth, J. Hartwigsen, R. Bhushan, J. Nicholls, and S. Jensen.** 2002. Clinical trial of E1B-deleted adenovirus (dl1520) gene therapy for hepatocellular carcinoma. Cancer Gene Ther **9:**254-9.

52. **Hackett, N. R., and R. G. Crystal.** 2004. Adenovirus Vectors for Gene Therapy, p. 17-41. *In* N. S. Templeton (ed.), Gene and Cell Therapy, 2nd ed. Marcel Dekker, New York.

53. **Hogle, J. M., M. Chow, and D. J. Filman.** 1985. Three-dimensional structure of poliovirus at 2.9 A resolution. Science **229:**1358-65.

54. **Hong, S. S., and P. Boulanger.** 1995. Protein ligands of the human adenovirus type 2 outer capsid identified by biopanning of a phage-displayed peptide library on separate domains of wild-type and mutant penton capsomers. EMBO J. **14:**4714-27.

55. **Jiang, W., M. L. Baker, S. J. Ludtke, and W. Chiu.** 2001. Bridging the information gap: computational tools for intermediate resolution structure interpretation. J Mol Biol **308:**1033-44.

56. **Jiang, W., and S. J. Ludtke.** 2005. Electron cryomicroscopy of single particles at subnanometer resolution. Curr. Opin. Struct. Biol. **15:**571-7.

57. **Kauffman, R. S., and H. S. Ginsberg.** 1976. Characterization of a temperature-sensitive, hexon transport mutant of type 5 adenovirus. J Virol **19:**643-58.

58. **Kawashima, H., A. Ogose, T. Yoshizawa, R. Kuwano, Y. Hotta, T. Hotta, H. Hatano, H. Kawashima, and N. Endo.** 2003. Expression of the coxsackievirus and adenovirus receptor in musculoskeletal tumors and mesenchymal tissues: efficacy of adenoviral gene therapy for osteosarcoma. Cancer Sci **94:**70-5.

59. **Kielian, M., and F. A. Rey.** 2006. Virus membrane-fusion proteins: more than one way to make a hairpin. Nat Rev Microbiol **4:**67-76.

60. **Kirkegaard, K.** 1990. Mutations in VP1 of poliovirus specifically affect both encapsidation and release of viral RNA. J Virol **64:**195-206.

61. **Lehmberg, E., J. A. Traina, J. A. Chakel, R. J. Chang, M. Parkman, M. T. McCaman, P. K. Murakami, V. Lahidji, J. W. Nelson, W. S. Hancock, E. Nestaas, and E. Pungor, Jr.** 1999. Reversed-phase high-performance liquid chromatographic assay for the adenovirus type 5 proteome. J. Chromatogr. B Biomed. Sci. Appl. **732:**411-23.

62. **Leopold, P. L., G. Kreitzer, N. Miyazawa, S. Rempel, K. K. Pfister, E. Rodriguez-Boulan, and R. G. Crystal.** 2000. Dynein- and microtubule-mediated translocation of adenovirus serotype 5 occurs after endosomal lysis. Hum Gene Ther **11:**151-65.

63. **Lescar, J., A. Roussel, M. W. Wien, J. Navaza, S. D. Fuller, G. Wengler, G. Wengler, and F. A. Rey.** 2001. The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. Cell **105:**137-48.

64. **Liemann, S., K. Chandran, T. S. Baker, M. L. Nibert, and S. C. Harrison.** 2002. Structure of the reovirus membrane-penetration protein, Mu1, in a complex with is protector protein, Sigma3. Cell **108:**283-95.

65. **Liu, G. Q., L. E. Babiss, F. C. Volkert, C. S. Young, and H. S. Ginsberg.** 1985. A thermolabile mutant of adenovirus 5 resulting from a substitution mutation in the protein VIII gene. J Virol **53:**920-5.

66. **Ludtke, S. J., P. R. Baldwin, and W. Chiu.** 1999. EMAN: semiautomated software for high-resolution single-particle reconstructions. J. Struct. Biol. **128:**82-97.

67. **Ludtke, S. J., D. H. Chen, J. L. Song, D. T. Chuang, and W. Chiu.** 2004. Seeing GroEL at 6 A resolution by single particle electron cryomicroscopy. Structure **12:**1129-36.

68. **Lupas, A., M. Van Dyke, and J. Stock.** 1991. Predicting coiled coils from protein sequences. Science **252:**1162-4.

69. **Magnuson, B., E. K. Rainey, T. Benjamin, M. Baryshev, S. Mkrtchian, and B. Tsai.** 2005. ERp29 triggers a conformational change in polyomavirus to stimulate membrane binding. Mol Cell **20:**289-300.

70. **Makower, D., A. Rozenblit, H. Kaufman, M. Edelman, M. E. Lane, J. Zwiebel, H. Haynes, and S. Wadler.** 2003. Phase II clinical trial of intralesional administration of the oncolytic adenovirus ONYX-015 in patients with hepatobiliary tumors with correlative p53 studies. Clin Cancer Res **9:**693-702.

71. **Mangel, W. F., W. J. McGrath, D. L. Toledo, and C. W. Anderson.** 1993. Viral DNA and a viral peptide can act as cofactors of adenovirus virion proteinase activity. Nature **361:**274-5.

72. **Matthews, D. A., and W. C. Russell.** 1995. Adenovirus protein-protein interactions: molecular parameters governing the binding of protein VI to hexon and the activation of the adenovirus 23K protease. J Gen Virol **76 (Pt 8):**1959-69.

73. **McGrath, W. J., J. Ding, A. Didwania, R. M. Sweet, and W. F. Mangel.** 2003. Crystallographic structure at 1.6-A resolution of the human adenovirus proteinase

in a covalent complex with its 11-amino-acid peptide cofactor: insights on a new fold. Biochim Biophys Acta **1648:**1-11.

74. **Mendelsohn, C. L., E. Wimmer, and V. R. Racaniello.** 1989. Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. Cell **56:**855-65.

75. **Miller, C. R., D. J. Buchsbaum, P. N. Reynolds, J. T. Douglas, G. Y. Gillespie, M. S. Mayo, D. Raben, and D. T. Curiel.** 1998. Differential susceptibility of primary and established human glioma cells to adenovirus infection: targeting via the epidermal growth factor receptor achieves fiber receptor-independent gene transfer. Cancer Res **58:**5738-48.

76. **Mindell, J. A., and N. Grigorieff.** 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. J. Struct. Biol. **142:**334-47.

77. **Modis, Y., S. Ogata, D. Clements, and S. C. Harrison.** 2003. A ligand-binding pocket in the dengue virus envelope glycoprotein. Proc Natl Acad Sci U S A **100:**6986-91.

78. **Modis, Y., S. Ogata, D. Clements, and S. C. Harrison.** 2004. Structure of the dengue virus envelope protein after membrane fusion. Nature **427:**313-9.

79. **Modis, Y., S. Ogata, D. Clements, and S. C. Harrison.** 2005. Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. J Virol **79:**1223-31.

80. **Moscufo, N., A. G. Yafal, A. Rogove, J. Hogle, and M. Chow.** 1993. A mutation in VP4 defines a new step in the late stages of cell entry by poliovirus. J Virol **67:**5075-8.

81. **Nakano, M. Y., K. Boucke, M. Suomalainen, R. P. Stidwill, and U. F. Greber.** 2000. The first step of adenovirus type 2 disassembly occurs at the cell surface, independently of endocytosis and escape to the cytosol. J. Virol. **74:**7085-95.

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

83. **Ni, S., A. Gaggar, N. Di Paolo, Z. Y. Li, Y. Liu, R. Strauss, P. Sova, J. Morihara, Q. Feng, N. Kiviat, P. Toure, P. S. Sow, and A. Lieber.** 2006. Evaluation of adenovirus vectors containing serotype 35 fibers for tumor targeting. Cancer Gene Ther. **in press**

84. **Nibert, M. L., A. L. Odegard, M. A. Agosto, K. Chandran, and L. A. Schiff.** 2005. Putative autocleavage of reovirus mu1 protein in concert with outer-capsid disassembly and activation for membrane permeabilization. J Mol Biol **345:**461-74.

85. **Nibert, M. L., L. A. Schiff, and B. N. Fields.** 1991. Mammalian reoviruses contain a myristoylated structural protein. J Virol **65:**1960-7.

86. **Odegard, A. L., K. Chandran, X. Zhang, J. S. Parker, T. S. Baker, and M. L. Nibert.** 2004. Putative autocleavage of outer capsid protein micro1, allowing release of myristoylated peptide micro1N during particle uncoating, is critical for cell entry by reovirus. J Virol **78:**8732-45.

87. **Ophorst, O. J., S. Kostense, J. Goudsmit, R. L. De Swart, S. Verhaagh, A. Zakhartchouk, M. Van Meijer, M. Sprangers, G. Van Amerongen, S. Yuksel, A. D. Osterhaus, and M. J. Havenga.** 2004. An adenoviral type 5 vector carrying a type 35 fiber as a vaccine vehicle: DC targeting, cross neutralization, and immunogenicity. Vaccine **22:**3035-44.

88. **Orlova, E. V., P. Dube, J. R. Harris, E. Beckman, F. Zemlin, J. Markl, and M. van Heel.** 1997. Structure of keyhole limpet hemocyanin type 1 (KLH1) at 15 A resolution by electron cryomicroscopy and angular reconstitution. J Mol Biol **271:**417-37.

89. **Parashar, U. D., E. G. Hummelman, J. S. Bresee, M. A. Miller, and R. I. Glass.** 2003. Global illness and deaths caused by rotavirus disease in children. Emerg Infect Dis **9:**565-72.

90. **Parks, R. J.** 2005. Adenovirus protein IX: a new look at an old protein. Mol. Ther. **11:**19-25.

91. **Pesavento, J. B., S. E. Crawford, E. Roberts, M. K. Estes, and B. V. Prasad.** 2005. pH-induced conformational change of the rotavirus VP4 spike: implications for cell entry and antibody neutralization. J Virol **79:**8572-80.

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

93. **Rancourt, C., H. Keyvani-Amineh, S. Sircar, P. Labrecque, and J. M. Weber.** 1995. Proline 137 is critical for adenovirus protease encapsidation and activation but not enzyme activity. Virology **209:**167-73.

94. **Raper, S. E., N. Chirmule, F. S. Lee, N. A. Wivel, A. Bagg, G. P. Gao, J. M. Wilson, and M. L. Batshaw.** 2003. Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. Mol Genet Metab **80:**148-58.

95. **Reinisch, K. M., M. L. Nibert, and S. C. Harrison.** 2000. Structure of the reovirus core at 3.6 A resolution. Nature **404:**960-7.

96. **Rexroad, J., C. M. Wiethoff, A. P. Green, T. D. Kierstead, M. O. Scott, and C. R. Middaugh.** 2003. Structural stability of adenovirus type 5. J. Pharm. Sci. **92:**665-78.

97. **Rey, F. A., F. X. Heinz, C. Mandl, C. Kunz, and S. C. Harrison.** 1995. The envelope glycoprotein from tick-borne encephalitis virus at 2 A resolution. Nature **375:**291-8.

98. **Rosa-Calatrava, M., L. Grave, F. Puvion-Dutilleul, B. Chatton, and C. Kedinger.** 2001. Functional analysis of adenovirus protein IX identifies domains involved in capsid stability, transcriptional activity, and nuclear reorganization. J. Virol. **75:**7131-41.

99. **Rosa-Calatrava, M., F. Puvion-Dutilleul, P. Lutz, D. Dreyer, H. de The, B. Chatton, and C. Kedinger.** 2003. Adenovirus protein IX sequesters host-cell promyelocytic leukaemia protein and contributes to efficient viral proliferation. EMBO Rep **4:**969-75.

100. **Roussel, A., J. Lescar, M. C. Vaney, G. Wengler, G. Wengler, and F. A. Rey.** 2006. Structure and interactions at the viral surface of the envelope protein E1 of Semliki Forest virus. Structure **14:**75-86.

101. **Ruigrok, R. W., A. Barge, C. Albiges-Rizo, and S. Dayan.** 1990. Structure of adenovirus fibre. II. Morphology of single fibres. J Mol Biol **215:**589-96.

102. **Russell, W. C., and B. Precious.** 1982. Nucleic acid-binding properties of adenovirus structural polypeptides. J. Gen. Virol. **63:**69-79.

103. **Rux, J. J., P. R. Kuser, and R. M. Burnett.** 2003. Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution x-ray crystallographic, molecular modeling, and sequence-based methods. J. Virol. **77:**9553-66.

104. **Saban, S. D., R. R. Nepomuceno, L. D. Gritton, G. R. Nemerow, and P. L. Stewart.** 2005. CryoEM structure at 9A resolution of an adenovirus vector targeted to hematopoietic cells. J. Mol. Biol. **349:**526-37.

105. **Saban, S. D., M. Silvestry, G. R. Nemerow, and P. L. Stewart.** 2006. Visualization of {alpha}-helices in a 6 Angstrom resolution cryoEM structure of adenovirus allows refinement of capsid protein assignments. J Virol. **in press**

106. **Sachs, M. D., K. A. Rauen, M. Ramamurthy, J. L. Dodson, A. M. De Marzo, M. J. Putzi, M. P. Schoenberg, and R. Rodriguez.** 2002. Integrin alpha(v) and coxsackie adenovirus receptor expression in clinical bladder cancer. Urology **60:**531-6.

107. **Sahli, R., R. Freund, T. Dubensky, R. Garcea, R. Bronson, and T. Benjamin.** 1993. Defect in entry and altered pathogenicity of a polyoma virus mutant blocked in VP2 myristylation. Virology **192:**142-53.

108. **Schoehn, G., P. Fender, J. Chroboczek, and E. A. Hewat.** 1996. Adenovirus 3 penton dodecahedron exhibits structural changes of the base on fibre binding. EMBO J. **15:**6841-6.

109. **Segerman, A., J. P. Atkinson, M. Marttila, V. Dennerquist, G. Wadell, and N. Arnberg.** 2003. Adenovirus type 11 uses CD46 as a cellular receptor. J. Virol. **77:**9183-91.

110. **Seth, P., M. C. Willingham, and I. Pastan.** 1984. Adenovirus-dependent release of 51Cr from KB cells at an acidic pH. J Biol Chem **259:**14350-3.

111. **Seth, P., M. C. Willingham, and I. Pastan.** 1985. Binding of adenovirus and its external proteins to Triton X-114. Dependence on pH. J Biol Chem **260:**14431-4.

112. **Shah, A. K., and P. L. Stewart.** 1998. QVIEW: software for rapid selection of particles from digital electron micrographs. J. Struct. Biol. **123:**17-21.

113. **Shaw, A. L., R. Rothnagel, D. Chen, R. F. Ramig, W. Chiu, and B. V. Prasad.** 1993. Three-dimensional visualization of the rotavirus hemagglutinin structure. Cell **74:**693-701.

114. **Shayakhmetov, D. M., C. A. Carlson, H. Stecher, Q. Li, G. Stamatoyannopoulos, and A. Lieber.** 2002. A high-capacity, capsid-modified hybrid adenovirus/adeno-associated virus vector for stable transduction of human hematopoietic cells. J. Virol. **76:**1135-43.

115. **Shayakhmetov, D. M., and A. Lieber.** 2000. Dependence of adenovirus infectivity on length of the fiber shaft domain. J. Virol. **74:**10274-86.

116. **Shenk, T.** 1996. *Adenoviridae:* The Viruses and Their Replication, p. 2111-2148. *In* B. N. Fields, Knipe, D.M., Howley, P.M. (ed.), *Fields Virology*, vol. 2. Lippincott - Raven, Philadelphia.

117. **Smith, T. A., N. Idamakanti, M. L. Rollence, J. Marshall-Neff, J. Kim, K. Mulgrew, G. R. Nemerow, M. Kaleko, and S. C. Stevenson.** 2003. Adenovirus serotype 5 fiber shaft influences in vivo gene transfer in mice. Hum. Gene Ther. **14:**777-87.

118. **St George, J. A.** 2003. Gene therapy progress and prospects: adenoviral vectors. Gene Ther. **10:**1135-41.

119. **Stehle, T., S. J. Gamblin, Y. Yan, and S. C. Harrison.** 1996. The structure of simian virus 40 refined at 3.1 A resolution. Structure **4:**165-82.

120. **Stehle, T., Y. Yan, T. L. Benjamin, and S. C. Harrison.** 1994. Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. Nature **369:**160-3.

121. **Stewart, P. L., R. B. Cary, S. R. Peterson, and C. Y. Chiu.** 2000. Digitally collected cryo-electron micrographs for single particle reconstruction. Microsc Res Tech **49:**224-32.

122. **Stewart, P. L., T. S. Dermody, and G. R. Nemerow.** 2003. Structural basis of nonenveloped virus cell entry. Adv. Protein Chem. **64:**455-491.

123. **Stewart, P. L., S. D. Fuller, and R. M. Burnett.** 1993. Difference imaging of adenovirus: bridging the resolution gap between X-ray crystallography and electron microscopy. EMBO J. **12:**2589-99.

124. **Suloway, C., J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C. S. Potter, and B. Carragher.** 2005. Automated molecular microscopy: the new Leginon system. J. Struct. Biol. **151:**41-60.

125. **Takahashi, E., S. L. Cohen, P. K. Tsai, and J. A. Sweeney.** 2006. Quantitation of adenovirus type 5 empty capsids. Anal Biochem **349:**208-17.

126. **Tarbouriech, N., J. Curran, R. W. Ruigrok, and W. P. Burmeister.** 2000. Tetrameric coiled coil domain of Sendai virus phosphoprotein. Nat Struct Biol **7:**777-81.

127. **Tollefson, A. E., A. Scaria, T. W. Hermiston, J. S. Ryerse, L. J. Wold, and W. S. Wold.** 1996. The adenovirus death protein (E3-11.6K) is required at very late stages of infection for efficient cell lysis and release of adenovirus from infected cells. J Virol **70:**2296-306.

128. **Tosteson, M. T., and M. Chow.** 1997. Characterization of the ion channels formed by poliovirus in planar lipid membranes. J Virol **71:**507-11.

129. **Tosteson, M. T., M. L. Nibert, and B. N. Fields.** 1993. Ion channels induced in lipid bilayers by subvirion particles of the nonenveloped mammalian reoviruses. Proc Natl Acad Sci U S A **90:**10549-52.

130. **Van Heel, M.** 1987. Similarity measures between images. Ultramicroscopy **21:**95-100.

131. **van Heel, M., G. Harauz, E. V. Orlova, R. Schmidt, and M. Schatz.** 1996. A new generation of the IMAGIC image processing system. J. Struct. Biol. **116:**17-24.

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

133. **van Raaij, M. J., A. Mitraki, G. Lavigne, and S. Cusack.** 1999. A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. Nature **401:**935-8.

134. **Varghese, R., Y. Mikyas, P. L. Stewart, and R. Ralston.** 2004. Postentry neutralization of adenovirus type 5 by an antihexon antibody. J. Virol. **78:**12320-32.

135. **Vellekamp, G., F. W. Porter, S. Sutjipto, C. Cutler, L. Bondoc, Y. H. Liu, D. Wylie, S. Cannon-Carlson, J. T. Tang, A. Frei, M. Voloch, and S. Zhuang.** 2001. Empty capsids in column-purified recombinant adenovirus preparations. Hum Gene Ther **12:**1923-36.

136. **Vihinen-Ranta, M., D. Wang, W. S. Weichert, and C. R. Parrish.** 2002. The VP1 N-terminal sequence of canine parvovirus affects nuclear transport of capsids and efficient cell infection. J Virol **76:**1884-91.

137. **Weber, J.** 1976. Genetic analysis of adenovirus type 2 III. Temperature sensitivity of processing viral proteins. J Virol **17:**462-71.

138. **Webster, A., R. T. Hay, and G. Kemp.** 1993. The adenovirus protease is activated by a virus-coded disulphide-linked peptide. Cell **72:**97-104.

139. **Wickham, T. J., P. Mathias, D. A. Cheresh, and G. R. Nemerow.** 1993. Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. Cell **73:**309-19.

140. **Wiethoff, C. M., H. Wodrich, L. Gerace, and G. R. Nemerow.** 2005. Adenovirus protein VI mediates membrane disruption following capsid disassembly. J Virol **79:**1992-2000.

141. **Wodrich, H., T. Guan, G. Cingolani, D. Von Seggern, G. Nemerow, and L. Gerace.** 2003. Switch from capsid protein import to adenovirus assembly by cleavage of nuclear transport signals. EMBO J. **22:**6245-55.

142. **Wolf, E., P. S. Kim, and B. Berger.** 1997. MultiCoil: a program for predicting two- and three-stranded coiled coils. Protein Sci. **6:**1179-89.

143. **Wolf, M., D. J. DeRosier, and N. Grigorieff.** 2006 Ewald sphere correction for single-particle electron microscopy. Ultramicroscopy **106(4-5):**376-82.

144. **Wu, E., and G. R. Nemerow.** 2004. Virus yoga: the role of flexibility in virus host cell recognition. Trends Microbiol. **12:**162-9.

145. **Wu, E., L. Pache, D. J. Von Seggern, T. M. Mullen, Y. Mikyas, P. L. Stewart, and G. R. Nemerow.** 2003. Flexibility of the adenovirus fiber is required for efficient receptor interaction. J. Virol. **77:**7225-35.

146. **Wu, E., S. A. Trauger, L. Pache, T. M. Mullen, D. J. von Seggern, G. Siuzdak, and G. R. Nemerow.** 2004. Membrane cofactor protein is a receptor for adenoviruses associated with epidemic keratoconjunctivitis. J. Virol. **78:**3897-905.

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

148. **Yeager, M., J. A. Berriman, T. S. Baker, and A. R. Bellamy.** 1994. Three-dimensional structure of the rotavirus haemagglutinin VP4 by cryo-electron microscopy and difference map analysis. Embo J **13:**1011-8.

149. **Yoder, J. D., and P. R. Dormitzer.** 2006. Alternative intermolecular contacts underlie the rotavirus VP5* two- to three-fold rearrangement. Embo J **25:**1559-68.

150. **Yonekura, K., S. Maki-Yonekura, and K. Namba.** 2003. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. Nature **424:**643-50.

151. **You, Z., D. C. Fischer, X. Tong, A. Hasenburg, E. Aguilar-Cordova, and D. G. Kieback.** 2001. Coxsackievirus-adenovirus receptor expression in ovarian cancer cell lines is associated with increased adenovirus transduction efficiency and transgene expression. Cancer Gene Ther **8:**168-75.

152. **Zadori, Z., J. Szelei, M. C. Lacoste, Y. Li, S. Gariepy, P. Raymond, M. Allaire, I. R. Nabi, and P. Tijssen.** 2001. A viral phospholipase A2 is required for parvovirus infectivity. Dev Cell **1:**291-302.

153. **Zhang, X., Y. Ji, L. Zhang, S. C. Harrison, D. C. Marinescu, M. L. Nibert, and T. S. Baker.** 2005. Features of reovirus outer capsid protein mu1 revealed by electron cryomicroscopy and image reconstruction of the virion at 7.0 Angstrom resolution. Structure **13:**1545-57.

154. **Zhang, Y., W. Zhang, S. Ogata, D. Clements, J. H. Strauss, T. S. Baker, R. J. Kuhn, and M. G. Rossmann.** 2004. Conformational changes of the flavivirus E glycoprotein. Structure **12:**1607-18.

155. **Zhou, Z. H., and W. Chiu.** 2003. Determination of icosahedral virus structures by electron cryomicroscopy at subnanometer resolution. Adv. Protein Chem. **64:**93-124.

156. **Zhou, Z. H., M. Dougherty, J. Jakana, J. He, F. J. Rixon, and W. Chiu.** 2000. Seeing the herpesvirus capsid at 8.5 A. Science **288:**877-80.

157. **Zubieta, C., G. Schoehn, J. Chroboczek, and S. Cusack.** 2005. The structure of the human adenovirus 2 penton. Mol. Cell. **17:**121-35.