Adenoviral Vectors for Gene Therapy

Second Edition

Edited by

David T. Curiel

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE• SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, UK
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

First Edition 2002

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-800276-6

Acquisition Editor: Linda Versteeg-Buschman
Editorial Project Manager: Halima Williams
Production Project Manager: Karen East and Kirsty Halterman
Designer: Alan Studholme

Typeset by TNQ Books and Journals
www.tnq.co.in

List of Contributors

Yadvinder S. Ahi HIV Drug Resistance Program, National Cancer Institute, Frederick National Laboratory for Cancer Research, Frederick, MD, USA

Steven M. Albelda Thoracic Oncology Research Group, Pulmonary, Allergy, and Critical Care Division, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Yasser A. Aldhamen Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI, USA

Ramon Alemany IDIBELL-Institut Català d'Oncologia, L'Hospitalet de Llobregat, Barcelona, Spain

Marta M. Alonso Department of Medical Oncology, Clínica Universidad de Navarra, University of Navarra, Pamplona, Spain

P.M. Alves iBET, Instituto de Biologia Experimental e Tecnológica, Oeiras, Portugal; Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

Andrea Amalfitano Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI, USA; College of Osteopathic Medicine, Michigan State University, East Lansing, MI, USA

Rachael Anatol Office of Cellular, Tissue, and Gene Therapies, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

C.A. Anderson Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

Svetlana Atasheva Lowance Center for Human Immunology, Departments of Pediatrics and Medicine, Emory University, Atlanta, GA, USA

Michael A. Barry Division of Infectious Diseases, Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA; Department of Immunology, Mayo Clinic, Rochester, MN, USA; Department of Molecular Medicine, Mayo Clinic, Rochester, MN, USA

Raj K. Batra UCLA School of Medicine, Division of Pulmonary and Critical Care Medicine, GLA-VAHCS, Los Angeles, CA, USA; Jonsson Comprehensive Cancer Center, UCLA, Los Angeles, CA, USA

A.J. Bett Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

A. Bout Crucell NV, Leiden, The Netherlands

K. Brouwer Crucell NV, Leiden, The Netherlands

Nicola Brunetti-Pierri Telethon Institute of Genetics and Medicine, Pozzuoli, Italy; Department of Translational Medicine, Federico II University, Naples, Italy

Andrew P. Byrnes Division of Cellular and Gene Therapies, FDA Center for Biologics Evaluation and Research, Silver Spring, MD, USA

Shyambabu Chaurasiya Department of Oncology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

L. Chen Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

A.S. Coroadinha iBET, Instituto de Biologia Experimental e Tecnológica, Oeiras, Portugal; Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

Igor P. Dmitriev Department of Radiation Oncology, School of Medicine, Washington University, St. Louis, MO, USA

Hildegund C.J. Ertl Wistar Institute, Philadelphia, PA, USA

P. Fernandes iBET, Instituto de Biologia Experimental e Tecnológica, Oeiras, Portugal; Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal; Autolus, London, UK

Juan Fueyo Department of Neuro-Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA; Department of Neurosurgery, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

S.M. Galloway Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

Thomas A. Gardner Department of Urology, Indiana University Medical Center, Indianapolis, IN, USA; Department of Microbiology and Immunology, Indiana University Medical Center, Indianapolis, IN, USA

Candelaria Gomez-Manzano Department of Neuro-Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA; Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

Urs F. Greber Institute of Molecular Life Sciences, University of Zurich, Zurich, Switzerland

Diana Guimet Department of Molecular Genetics and Microbiology, School of Medicine, Stony Brook University, Stony Brook, NY, USA

Michael Havert Office of Cellular, Tissue, and Gene Therapies, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

Patrick Hearing Department of Molecular Genetics and Microbiology, School of Medicine, Stony Brook University, Stony Brook, NY, USA

Masahisa Hemmi Laboratory of Biochemistry and Molecular Biology, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka, Japan

R.B. Hill Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

Mary M. Hitt Department of Oncology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

Ying Huang Office of Cellular, Tissue, and Gene Therapies, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

Ilan Irony Office of Cellular, Tissue, and Gene Therapies, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

Hong Jiang Department of Neuro-Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

Sergey A. Kaliberov Department of Radiation Oncology, School of Medicine, Washington University, St. Louis, MO, USA

Chinghai H. Kao Department of Urology, Indiana University Medical Center, Indianapolis, IN, USA; Department of Microbiology and Immunology, Indiana University Medical Center, Indianapolis, IN, USA

Dayananda Kasala Department of Bioengineering, College of Engineering, Hanyang University, Seongdong-gu, Seoul, Republic of Korea

D. Kaslow Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

Benjamin B. Kasten Department of Radiology, The University of Alabama at Birmingham, Birmingham, AL, USA

Johanna K. Kaufmann German Cancer Research Center (DKFZ), Heidelberg, Germany

Jay K. Kolls Richard King Mellon Foundation Institute for Pediatric Research, Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA, USA; Department of Pediatrics, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA

Johanna P. Laakkonen Department of Biotechnology and Molecular Medicine, A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio, Finland

R. Lardenoije Crucell NV, Leiden, The Netherlands

J. Lebron Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

B.J. Ledwith Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

J. Lewis Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

Erik Lubberts Department of Immunology, Erasmus MC, University Medical Center, Rotterdam, The Netherlands; Department of Rheumatology, Erasmus MC, University Medical Center, Rotterdam, The Netherlands

Stefania Luisoni Institute of Molecular Life Sciences, University of Zurich, Zurich, Switzerland

S.V. Machotka Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

S. Manam Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

D. Martinez Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

Suresh K. Mittal Department of Comparative Pathobiology, College of Veterinary Medicine and Purdue University Center for Cancer Research, Purdue University, West Lafayette, IN, USA

Hiroyuki Mizuguchi Laboratory of Biochemistry and Molecular Biology, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka, Japan

Edmund Moon Thoracic Oncology Research Group, Pulmonary, Allergy, and Critical Care Division, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Stephen J. Murphy Molecular Medicine Program, Mayo Clinic, Rochester, MN, USA

Dirk M. Nettelbeck German Cancer Research Center (DKFZ), Heidelberg, Germany

Philip Ng Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

W.W. Nichols Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

Raymond John Pickles Cystic Fibrosis/Pulmonary Research and Treatment Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Sudhanshu P. Raikwar Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of Missouri and Harry S. Truman Veterans' Memorial Hospital, Columbia, MO, USA

Paul N. Reynolds Department of Thoracic Medicine and Lung Research Laboratory, Royal Adelaide Hospital, Adelaide

Jillian R. Richter Department of Radiology, The University of Alabama at Birmingham, Birmingham, AL, USA

Yisel Rivera-Molina Department of Neuro-Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

Qian Ruan PaxVax Inc., San Diego, CA, USA

C. Russo Merck Research Laboratories, Merck & Co., Inc., West Point, PA, USA

Carl Scandella Carl Scandella Consulting, Bellevue, WA, USA

Paul Shabram PaxVax Inc., San Diego, CA, USA

Anurag Sharma Department of Pediatrics, Weill Cornell Medical College, New York, NY, USA

Sherven Sharma UCLA/Wadsworth Pulmonary Immunology Laboratory, Division of Pulmonary and Critical Care Medicine, GLA-VAHCS, Los Angeles, CA, USA

Dmitry M. Shayakhmetov Lowance Center for Human Immunology, Departments of Pediatrics and Medicine, Emory University, Atlanta, GA, USA

A.C. Silva iBET, Instituto de Biologia Experimental e Tecnológica, Oeiras, Portugal; Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

Phoebe L. Stewart Department of Pharmacology and Cleveland Center for Membrane and Structural Biology, Case Western Reserve University, Cleveland, OH, USA

Hideyo Ugai Cancer Biology Division, Department of Radiation Oncology, School of Medicine, Washington University, St. Louis, MO, USA

D. Valerio Crucell NV, Leiden, The Netherlands

M. van der Kaaden Crucell NV, Leiden, The Netherlands

Gary Vellekamp Vellekamp Consulting LLC, Montclair, NJ, USA

Sai V. Vemula Laboratory of Molecular Virology, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

Richard G. Vile Molecular Medicine Program, Mayo Clinic, Rochester, MN, USA

R. Vogels Crucell NV, Leiden, The Netherlands

Stefan Worgall Department of Pediatrics, Weill Cornell Medical College, New York, NY, USA; Department of Genetic Medicine, Weill Cornell Medical College, New York, NY, USA

Lily Wu Jonsson Comprehensive Cancer Center, UCLA, Los Angeles, CA, USA; Department of Urology, UCLA School of Medicine, Los Angeles, CA, USA; Department of Pediatrics, UCLA School of Medicine, Los Angeles, CA, USA

Enric Xipell Department of Medical Oncology, Clínica Universidad de Navarra, University of Navarra, Pamplona, Spain

Seppo Ylä-Herttuala Department of Biotechnology and Molecular Medicine, A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio, Finland; Department of Medicine, University of Eastern Finland, Kuopio, Finland; Gene Therapy Unit, Kuopio University Hospital, Kuopio, Finland

Chae-Ok Yun Department of Bioengineering, College of Engineering, Hanyang University, Seongdong-gu, Seoul, Republic of Korea

Kurt R. Zinn Department of Radiology, The University of Alabama at Birmingham, Birmingham, AL, USA

D. Zuidgeest Crucell NV, Leiden, The Netherlands

Adenovirus Structure

1

Phoebe L. Stewart
Department of Pharmacology and Cleveland Center for Membrane and Structural Biology, Case Western Reserve University, Cleveland, OH, USA

1. Historical Perspective on Adenovirus Structure

The structure of the adenovirus virion is quite complex and our understanding of it has been evolving from before 1965. Early negative stain electron micrographs of adenovirus revealed an icosahedral capsid with 252 capsomers and long fibers protruding from the vertices.[1] Later these capsomers were identified as 240 hexons and 12 pentons, with the pentons at the fivefold vertices of the capsid. The pentons each have five neighboring capsomers and the hexons each have six neighboring capsomers. As the adenoviral molecular components were identified and their stoichoimetries characterized, it became apparent that the hexons and pentons were different proteins. The hexons are trimeric proteins and the pentons are formed by two proteins, a pentameric penton base and a trimeric fiber.[2] Subsequently, X-ray crystallography provided atomic resolution structures of hexon,[3] penton base,[4] fiber,[5,6] and adenovirus protease,[7] which is involved in virion maturation. In addition to the three major protein components of the capsid (hexon, penton base, and fiber), there are four minor capsid proteins (proteins IIIa, VI, VIII, IX).[8,9] The minor proteins are also referred to as cement proteins as they serve to stabilize the capsid. They also play important roles in the assembly, disassembly, and cell entry of the virus. Atomic resolution structures have not yet been determined for the minor proteins isolated from the adenovirus capsid. However, cryo-electron microscopy (cryoEM) has provided moderate structural information on the density of the minor proteins in the context of the virion.[10–13] In 2010, atomic resolution structures of adenovirus were determined by cryoEM and X-ray crystallography.[14,15] Despite these two atomic, or near atomic, resolution (3.5–3.6 Å) structures, controversies remained regarding the structure and assignment of the minor capsid proteins. In 2014, a refined crystal structure of adenovirus at 3.8 Å resolution revised the minor capsid protein structures and locations.[16]

The adenoviral genome is relatively large, with ~30–40 kb.[8] It is notable in that large deletions and insertions can be tolerated, a feature that contributes to the enduring popularity of adenovirus as a gene delivery vector.[17] Within the core of the virion there are five proteins associated with the double-stranded DNA genome (proteins V, VII, mu, IVa2, and terminal binding protein).[9] The structure of the genome and how it is packaged with its associated proteins in the core of the virion is not well understood. Early negative stain EM and ion etching studies suggested that the core is organized as 12 large spherical nucleoprotein assemblies, termed adenosomes.[18,19] However, cryoEM and crystallographic structures of adenovirus show that the core does not follow the strict icosahedral symmetry of the capsid.[14–16]

Adenoviral Vectors for Gene Therapy. http://dx.doi.org/10.1016/B978-0-12-800276-6.00001-2

Adenovirus was one of the first samples imaged during the development of the cryoEM technique[20] and was among the first set of viruses to have its structure determined by the cryoEM single particle reconstruction method.[21] Since then cryoEM structures have been determined for multiple types of adenovirus and adenovirus in complex with various host factors.[10–12,14,22–29] Docking of crystal structures of capsid proteins into the cryoEM density and difference imaging have been useful approaches for dissecting the complex nature of the capsid. An early example of difference imaging was applied in two dimensions to scanning transmission electron microscopy (STEM) images of the group-of-nine hexons and this work helped to elucidate the position of protein IX within the icosahedral facet.[30] Difference imaging in three dimensions led to an early tentative assignment for the positions of the minor capsid proteins within the capsid based on copy number and approximate mass.[13] As higher resolution cryoEM structures were determined, some of these initial assignments were revised.[10–12] Visualization of α-helices was achieved with a 6 Å resolution cryoEM structure.[12] This structure facilitated more accurate docking of hexon and penton base crystal structures and produced a clearer difference map and more detailed density for the minor capsid proteins. Secondary structure prediction for the minor capsid proteins was used to tentatively assign density regions to minor capsid proteins. Determination of an atomic resolution (3.6 Å) structure by cryoEM was facilitated by the use of a high-end FEI Titan Krios electron microscope.[14] Micrographs for this dataset were collected on film and scanned for digital image processing. The final dataset included 31,815 individual particle images. The resolution was estimated by reference-based Fourier shell correlation coefficient and supported by observation of both α-helical and β-strand density. Density was also observed for some of the side chains, particularly bulky amino acids. The assignments for the minor capsid protein locations were assumed to be the same as interpreted from the 6 Å resolution cryoEM structure.[12] Atomic models were produced for minor capsid proteins IIIa, VIII, and IX from the atomic resolution cryoEM density map using bulky amino acids as landmarks.[14]

Attempts to crystallize intact adenovirus began in 1999 and proceeded for more than 10 years before the first atomic resolution crystal structure was published.[15,31] Several factors hampered early crystallization efforts, including the long protruding fiber, the instability of virions at certain pH values, the tendency of adenovirus particles to aggregate, and relatively low yields from standard virus preparations. Use of a vector based on human adenovirus type 5 (HAdV5), but with the short fiber from type 35 (Ad5.F35, also called Ad35F), helped to solve some of the production and crystallization difficulties. This vector was also used for several moderate resolution cryoEM structural studies.[11,12] Collection of diffraction data for atomic resolution structure determination spanned several years. Even though crystals were flash-cooled in liquid nitrogen, they were still highly radiation sensitive and only 2–5% of the crystals diffracted to high resolution. Diffraction data from nearly 900 crystals were collected but only a small subset of these data was used to generate the dataset. The best crystals diffracted well to 4.5 Å resolution and weakly to 3.5 Å at synchrotron sources. The initial phase information was derived from a pseudo-atomic capsid of adenovirus generated from fitting the crystallographic structures of hexon and penton base into a cryoEM structure of Ad5.F35 at 9 Å resolution.[11] In 2010, partial atomic models were built for some of the minor capsid proteins.[15]

After collection of more diffraction data and additional refinement a refined crystal structure was published with more complete models for minor capsid proteins IIIa, VI, VIII, and IX and surprisingly for a portion of the core protein V.[16] To compensate for the relatively modest resolution (3.8 Å) of the structure, a method was devised to evaluate the reliability of assigned amino acid sequences to the experimental electron density. This gives credence to the latest assignments for the locations of the minor capsid proteins within the capsid. It is important to recognize that adenovirus is one of the largest biomolecular assemblies with an atomic resolution structure determined by X-ray crystallography (>98,000 nonhydrogen atoms used in refinement of the asymmetric unit). With an assembly of this size and complexity and with less than ideal resolution data, assigning the locations of the minor capsid proteins is quite a challenging task.

There are over 60 HAdV types categorized in seven species (human adenovirus A–G). Species D adenoviruses are the most numerous, many of which were identified during the AIDS epidemic.[32] AIDS patients and other immunocompromised patients are particularly susceptible to adenovirus. Adenovirus causes acute respiratory illness, epidemic keratoconjunctivitis, acute hemorrhagic cystitis, hepatitis, myocarditis, and gastroenteritis in humans. Adenoviruses have also been characterized from the five major classes of vertebrate species, mammals, birds, reptiles, amphibians, and fish.[33] Structural studies of human and animal adenoviruses have contributed to our understanding of the molecular complexity within the Adenoviridae family.

2. Hexon Structure and Capsid Packing

The icosahedral capsid of adenovirus is composed of 240 trimeric hexons and 12 pentameric penton bases at the vertices with associated fibers. For HAdV2, hexons account for the majority (>83%) of the protein mass in the capsid.[34] The first hexon crystal structure was that of HAdV2.[3] At that time in 1986, the hexon subunit was the longest polypeptide whose structure was determined by X-ray crystallography with 967 residues per hexon monomer. Higher resolution (2.2 and 2.5 Å) crystal structures of HAdV2 and HAdV5 hexons are now available (PDB-ID: 1P2Z; PDB-ID: 1P30).[34] The hexon crystal structure revealed that although it is a trimeric protein, the base of the molecule is shaped as a hexagon, which is optimal for close packing within the capsid. The hexagonal base of the hexon trimer is formed by two viral jellyroll domains in each hexon monomer, with each jellyroll situated at a point of the hexagon. The topology of the jellyrolls is similar to that of icosahedral RNA viruses, although the architectural roles of the jellyrolls in forming the icosahedral capsids of these viruses are different.[3] Intriguingly, the hexon fold is the same as that of the major capsid protein P3 of the bacteriophage PRD1.[35]

The top of the hexon trimer is trimeric in shape with three protruding towers. Each tower is formed by intertwined loops from all three hexon monomers. The intertwining within the hexon trimer is so extensive that an accessory protein, called the 100k protein, is required to help fold the hexon trimer.[36,37] Hexon has a large subunit interface and each subunit of hexon clasps its neighboring subunit, resulting in a highly stable trimeric structure.[38]

Comparison of hexon sequences from multiple adenovirus types led to the finding of multiple hypervariable regions within the hexon.[39] Originally it was thought that only some of these regions were mapped to the top of the hexon. However, determination of the HAdV2 crystal structure at 2.5 Å resolution led to an atomic model with 25% of the sequence reassigned compared to the earlier HAdV5 crystal structure.[38] Later both the HAdV2 and the HAdV5 crystal structures were refined with newer protocols.[34] The hexons from these two adenovirus types are highly homologous (86% identity) and their refined structures are very similar. The revised HAdV2 and HAdV5 hexon crystal structures place all of the hypervariable loops near the exposed top of the hexon trimer.

Hexon sequences from different viral types also revealed a high level of sequence conservation within a particular human species (~88%), reduced conservation between types of different human species (79–81%), and less conservation between types of different animal species (66–68%).[39] The majority of the differences found in hexon sequences are within the hypervariable loops and these loops are often the targets of neutralizing antibodies.[40] Following vaccination and natural infection, neutralizing antibodies are produced to both hexon and fiber, although the response to hexon appears to be dominant.[41] The flexibility and sequence tolerance of the hexon hypervariable loops have made them useful as insertion sites for modification of the adenovirus capsid.[42]

On each hexon trimer between the three protruding towers that project from the outer viral surface is a central depression. CryoEM structures of adenovirus in complex with vitamin K-dependent blood coagulation factor X indicate that the hexon depression is the binding site for the GLA (γ-carboxyglutamic acid rich) domain of the factor.[23,26,29] Specifically, a single threonine residue (T425) of HAdV5 is critical for the interaction with factor X, as mutation of this residue in the context of the virion abrogates binding to factor X.[23] Injection of mice intravenously with this virus mutant indicated that it does not infect hepatocytes efficiently, whereas wild-type and other virus mutants with single or double hexon mutations are efficient in this regard. Factor X plays a role in mediating Ad-hepatocyte transduction in vivo after intravenous administration. The adenovirus/factor X complex utilizes an alternative cellular uptake pathway and the adenovirus-bound factor X interacts with heparan sulfate proteoglycan on macrophages.[23,43]

The hexons are arranged with 12 trimers in each of the 20 facets of the icosahedron. There are four unique positions for the hexon trimer within the asymmetric unit of the capsid (Figure 1). The asymmetric unit is the smallest repeating unit of the capsid and corresponds to one-third of an icosahedral facet. Although different conventions have been used for numbering the hexons, the most common convention labels the hexons next to the penton base as position 1, the hexons next to the icosahedral twofold axes as position 2, the hexons next to the icosahedral threefold axes as position 3, and the fourth remaining site as position 4 (Figure 1(A)). The hexons next to the penton base, which are also referred to as the peripentonal hexons, have been observed to dissociate separately from the other hexons.[44] The remaining hexons dissociate in groups-of-nine hexons. These nine hexons (three each in positions 2, 3, and 4) form the central part of each icosahedral facet. The group-of-nine hexons are held together by the minor capsid protein, protein IX.[30]

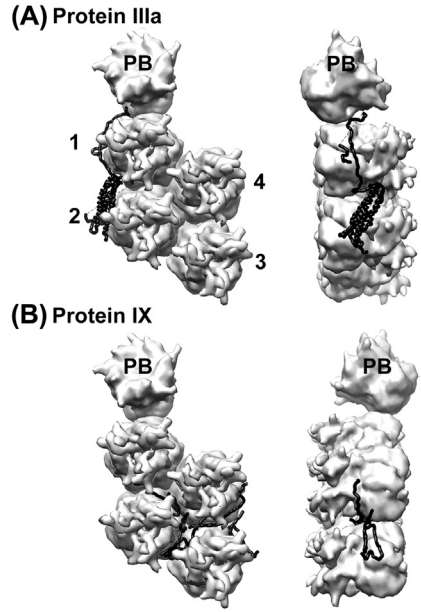

Figure 1 Structure and location of the outer capsid proteins as assigned in the refined adenovirus crystal structure.[16] (A) The enlarged asymmetric unit, with four independent hexon trimers (1–4) and a complete penton base (PB), is shown as a 5 Å surface representation (light gray) together with the ordered portion of protein IIIa (black). Protein IIIa is chain O in PDB-ID: 4CWU. (B) The enlarged asymmetric unit together with the ordered portions of four copies of protein IX (black). Only the N-terminal portions of protein IX are ordered. The four copies of protein IX in the asymmetric unit are chains P, Q, R, and S in PDB-ID: 4CWU. Top and side views are shown in both panels. Dashed lines represent disordered regions. This figure was made with UCSF Chimera.[126]

3. Penton Base Structure and Integrin-Binding RGD Loop

The penton base is a pentameric protein that is shaped as a pentagon and packs nicely at each vertex of the capsid within a ring of five peripentonal hexons. The penton base of human and animal adenovirus types is typically highly conserved with ~70% homology between the sequences of any two types.[4] Negative stain electron micrographs of the adenovirus penton, composed of penton base and fiber, showed a pentameric structure with the fiber shaft protruding from the center.[45] CryoEM structures of dodecahedra composed of 12 HAdV3 penton bases or complete pentons showed subtle changes in the penton base structure with fiber binding.[46]

The crystal structure of the penton base was first determined for an N-terminally truncated form of the HAdV2 protein that formed regular dodecahedral particles with 12 complete pentamers.[4] Two structures were determined at the same time, one of penton base alone and one with an N-terminal fragment of the fiber protein revealing how the fiber interacts with the penton base (PDB-ID: 1X9P; PDB-ID: 1X9T).[4] The crystal structure of the HAdV2 penton revealed that the top of the

penton base has grooves between the subunits that serve as binding sites for a conserved motif near the N-terminal end of fiber.[4] There is a symmetry mismatch between the trimeric fiber and the pentameric penton base, meaning that only three of the five grooves are occupied in each penton base of the assembled virion.[2] The pentameric form of the penton base buries a significant portion of the total surface area of each monomer. Mainly hydrophobic surfaces are buried in formation of the pentamer. The oligomeric penton base is composed of tilted monomers that form an assembly with an overall right-handed twist.

The pentagonal shape at the basal end of the molecule is formed by one jellyroll in each monomer. Intriguingly, the jellyroll within penton base is topologically related to the jellyroll in hexon. In addition to the jellyroll motif each monomer has an upper insertion domain, which protrudes from the outer capsid surface. The upper insertion domain is formed by two long insertion loops between strands of the jellyroll. One insertion loop contains the hypervariable Arg–Gly–Asp (RGD) region. This region is the most variable in sequence and length among adenovirus types. The RGD loop for HAdV2 is ~80 aa and is glycine- and alanine-rich. Most of the loop is flexible as no density is observed for residues 298–375 in the X-ray structure.[4] The second insertion, called the variable loop, forms a flexible β ribbon projecting from the top of penton base. In HAdV2 this loop is formed by residues 142–169, but in other adenovirus types it can be up to 10 residues longer.

The sequences of a penton base region including the RGD loop, variable loop, and surrounding residues from 51 human adenovirus types were used for phylogenetic analysis and structural prediction.[47] As expected, the phylogenetic analysis demonstrated clustering of the adenovirus types according to their species. In addition, clustering of the species B types supported the concept of dividing species B types into subspecies B1 and B2. Structural models for the various penton base proteins were built based on the crystallographic structure of the HAdV2 penton base. The divergence of the jellyroll motif compared to the HAdV2 penton base structure was predicted to be only 9.8–15.5%, whereas the divergence of the upper insertion domain was in the range of 37.3–38.8%.

Most, but not all, types of adenovirus have an RGD motif in one of the two surface loops of the penton base.[48] This motif is required for interactions with cellular integrins. Clustering of integrins on the host cell surface is promoted by interaction with penton base RGD loops and this leads to activation of signaling pathways that result in rapid internalization of the virus into clathrin-coated pits and endosomes.[49] The enteric adenovirus types HAdV40 and HAdV41 of species F lack the RGD motif on their penton base and do not utilize integrins for cell entry.[50,51]

Moderate resolution cryoEM structures have been determined for HAdV2 and HAdV12 in complex with soluble forms of αvβ5 integrin.[22,27] Modeling with integrin crystal structures indicates that only a maximum of four integrins can bind per penton base. This is consistent with the surface plasmon resonance measurement of 4.2 integrin molecules per HAdV2 penton base at close to saturation.[22] The spacing of the RGD protrusions on the penton base (~60 Å) appears to be too close to allow five integrin heterodimers to bind to one penton base. Modeling shows that there is room to bind four integrin heterodimers to one penton base, but significant flexibility within

the penton base RGD loops is required to accommodate this binding configuration. It was hypothesized that the strain arising from this symmetry-mismatched interaction might lead to a conformational change in the penton base and promote partial release of penton base pentamers from the capsid.[27]

The flexibility of the penton base RGD loops was first demonstrated by a cryoEM structure of HAdV2 in complex with a Fab fragment from a monoclonal antibody that binds a peptide region of penton base including RGD.[52] The HAdV2 penton base crystal structure is missing quite a large peptide region of 78 residues in the RGD loop due to disorder.[4] Alignment of penton base sequences from human adenovirus types indicates that HAdV12 has one of the shorter RGD loops, with just 15 residues corresponding to the missing 78 residues in HAdV2.[52] However, even the shorter HAdV12 RGD loop is flexible as indicated by the cryoEM structures of HAdV12 in complex with αvβ5 integrin.[22,27]

Submission of two penton base sequences, those of HAdV5 and HAdV19c, to the ProteinDisOrder System (PrDOS) prediction webserver indicated that these RGD loops are predicted to be intrinsically disordered.[24] The significance of having an intrinsically disordered RGD loop might be related to increasing the binding rate constant of penton base to integrins on the cell surface. It has been demonstrated that the binding of intrinsically disordered proteins to structured targets with strong electrostatic interactions enhances the binding rate constants by several orders of magnitude.[53]

The penton base RGD loops have been implicated in binding human alpha defensins, which are peptides of the innate immune system.[24] Human alpha defensin 5 (HD5) can inhibit cell entry of adenoviral types from species A, B1, B2, C, and E, whereas species D and F types are resistant.[28] CryoEM structures of adenovirus/defensin complexes led to a model in which the RGD loops of sensitive adenoviral types wrap around HD5 monomers or dimers at the interface between penton base and fiber and stabilize the penton base/fiber complex.[24,28] This stabilization effect is thought to prevent release of the adenoviral membrane lytic factor, protein VI, and therefore adenovirus cannot escape from the endosome and is degraded by the host cell in the lysosomal pathway.

4. Fiber Structure and Receptor Interactions

The fiber is composed of three distinct regions: a short penton base interaction region near the N terminus, a shaft domain with a variable number of repeats, and a distal knob domain, which interacts with various receptors. The first atomic resolution structural information for the fiber was for the knob domain of HAdV5 (PDB-ID: 1KNB).[6] The crystal structure revealed an eight-stranded antiparallel β-sandwich structure in each monomer. The trimeric knob has a large buried surface area, indicating that the trimer is probably the most prevalent form of the fiber in solution. Crystal structures have now been determined for fiber knobs of numerous human adenovirus types, including HAdV3, HAdV7, HAdV11, HAdV12, HAdV14, HAdV16, HAdV21, HAdV35, and HAdV37.[54–61] In addition, crystal structures have been determined for canine and porcine fiber knobs.[62,63] These structures all reveal the same overall fold for the knob domain.

Sequence alignment of the shaft domain of multiple adenovirus types showed a common 15-residue repeat pattern.[64] The fold of this repeat pattern was revealed in a crystal structure of the knob domain plus four repeat units of the shaft from the HAdV2 protein (PDB-ID: 1QIU).[5] The fiber shaft fold represents a new structural motif for fibrous proteins, named the triple β-spiral. This fold is characterized by an extended β-strand running parallel to the fiber axis, a turn with a conserved glycine or proline, a second β-strand, and a following solvent-exposed loop of variable length. This structural motif is also found in the shaft domain of the reovirus sigma-1 protein.[65]

The structure of a short peptide region near the N terminus of fiber, termed as the universal fiber motif, was revealed in the crystal structure of the HAdV2 penton base with a 21-residue fiber peptide (PDB-ID: 1X9T).[4] The universal fiber motif is a mostly hydrophobic peptide region (FNPVYPY) that binds at the top of penton base at the subunit interface. All of the interactions observed between the fiber peptide and the penton base involve the conserved motif of the fiber and highly conserved residues of the penton base with the exception of one residue (Lys-387 of HAdV2 penton base). This indicates that it is likely that there is a universal mode of association between the N-terminal fiber motifs and the penton bases of various adenovirus types. The interactions between the fiber N-terminal region and the penton base were confirmed in a model of the HAdV5 fiber built by homology modeling and fitting of models within a 3.6 Å resolution cryoEM structure of the intact HAdV5 virion.[66]

The fiber knob is responsible for interaction with a variety of host cell attachment receptors, including coxsackie-adenovirus receptor (CAR), CD46 (membrane cofactor protein), sialic acid-containing oligosaccharides, GD1a glycan, and desmoglein-2 (DSG-2).[67–69] Numerous crystal structures have been determined with fiber knobs of various adenoviral types in complex with CAR,[54,70] CD46,[56,71,72] and sialic acid-containing molecules.[55,63,67] Whereas CAR and CD46 bind on the side of the trimeric fiber knob, sialic acid for the most part binds at the top of the fiber knob near the threefold symmetry axis. One exception to this, however, is the structure of the canine adenovirus type 2 (CAdV2) knob in complex with sialic acid.[63] This structure shows a distinct binding site for sialic acid, still on the top of the knob but more toward the periphery. The observation that CAR and CD46 bind on the side of the adenoviral fiber knobs, while sialic acid binds on the top of the knobs from human adenoviruses, suggests that there may be situations in which one fiber binds two different attachment receptors. This possibility is supported by a crystal structure of the HAdV37 fiber knob in complex with both a CAR domain and sialyl-D-lactose.[63]

5. Atomic Resolution Cryo-Electron Microscopy and X-ray Crystallographic Adenovirus Structures

Of course to truly appreciate the structure of adenovirus it is necessary to obtain an atomic resolution structure of the intact virion. In 2010, atomic resolution structures were published as determined by both cryoEM[14] and X-ray crystallography.[15] Four years later in 2014, a refined crystal structure was published.[16] Interpretation of the

cryoEM structure was aided by the known structures of the major capsid proteins. In addition, the resolution was sufficient to observe density for bulky side chain, and de novo atomic models were created for several of the minor capsid proteins (PDB-ID: 3IYN). Solving of the crystal structure at 3.5 Å resolution was aided by a pseudo-atomic capsid produced by fitting the coordinates of isolated capsid proteins into a cryoEM density map.[11] The first crystal structure provides atomic descriptions of hexon and penton base together with partial models for some of the minor capsid proteins (PDB-ID: 1VSZ).

Both structures represent tremendous achievements given the large size of adenovirus, 150 MDa, and the complexity of the capsid with over 100,000 nonhydrogen atoms per asymmetric unit. The problem is that with this size and level of complexity, and with only partial side-chain densities apparent, the assignment of the minor capsid proteins is ambiguous and the two structures differ in their interpretations. The cryoEM structure is of HAdV5 and the crystal structures are of the HAdV5-based vector, Ad5.F35. In terms of molecular composition they should only vary in their fibers, with Ad5.F35 containing the shorter HAdV35 fiber. Given that the structure of penton base and a fiber fragment indicated a universal mode of association between fibers and penton bases of various adenovirus types,[4] it would seem to be a safe assumption that, except for the fibers and possible crystal packing effects, the structure of icosahedral capsid would be the same between HAdV5 and Ad5.F35.

A refined crystal structure at 3.8 Å resolution was published in 2014 with more complete atomic models for the minor capsid proteins (PCD-ID: 4CWU).[16] Ideally an atomic resolution crystal structure would be at high enough resolution to observe density for all, or most, of the side chains so that the assignment of density regions to specific amino acid sequences would be unambiguous. Unfortunately, this was not the case. Therefore a strategy was designed to evaluate and score assigned sequences to features in the experimental density map. This involved grouping the 20 amino acids into six groups based on side-chain size. Scores were assigned based on how well the sequence matched the density. Comparisons were made after shifting the amino acid sequence by one residue at a time. In addition, the N to C direction of each polypeptide chain was reversed and the scores recalculated to confirm that the best match for the density was chosen. This careful analysis of the X-ray density lends support to the assignments of the minor capsid proteins made by Reddy and Nemerow.[16]

6. Hexons in the Atomic Resolution Adenovirus Structures

Comparison of the hexon coordinates within the cryoEM and crystallographic atomic resolution adenovirus structures is complicated by the fact that the authors chose a different set of four unique hexons to include in the asymmetric unit, or basic repeating unit of the capsid.[14–16] The nomenclature of the four hexons is the same in all structures, with hexon 1 next to the penton base, hexon 2 next to the icosahedral twofold axes, hexon 3 next to the icosahedral threefold axes, and hexon 4 at the remaining position in the asymmetric unit. However, the four representative hexons of the cryoEM structure

were chosen to surround the four-helix bundle at a facet edge, while the four hexons of the X-ray structure are all on the same side of the four-helix bundle.

Within the crystal structure of Ad5.F35 all 12 of the independent hexon subunits have virtually identical folds with a ~1 Å root mean square deviation on superimposition.[15] The main differences between the hexon subunits within the Ad5.F35 crystal structure are found at the N- and C-termini. Both the cryoEM and the crystal structures of adenovirus report coordinates for a few extra residues at the N- and C-termini of hexon, compared to the crystal structure of the isolated HAdV5 hexon.[34] However, the details of the hexon N- and C-terminal tail structures differ somewhat. Both the cryoEM and the crystal structures provide coordinates for some of the residues in the hexon hypervariable loops, which were disordered in the isolated hexon structure.[34] When packed in the adenovirus capsid the hypervariable loops mediate interhexon interactions and interactions with other capsid proteins. Selection and superimposition of a matching set of four hexons from the full icosahedral capsids of both the cryoEM and the crystal structures reveal some differences in interpretation for the hexon hypervariable loop structures.

7. Conformational Differences of the Penton Base in the Atomic Resolution Adenovirus Structures

The crystal structure of the isolated HAdV2 penton base was determined with an N-terminal truncation missing the first 48 residues because the full-length protein was easily degraded.[4] The coordinates for the isolated HAdV2 penton base (PDB-ID: 1X9T) begin with residue 52. In the HAdV5 atomic resolution cryoEM structure additional residues (aa 37–51) are traced in the N-terminal tail of the penton base.[14] The HAdV2 and HAdV5 penton base proteins are highly homologous (98% identity) and the overall fold is nearly identical. In the cryoEM structure the N-terminal residues of the HAdV5 penton base are observed to interact with a minor capsid protein below the penton base and then turn inward to connect to the genomic core. However, the N-terminal extensions of the penton base are not identified in the crystal structure of Ad5.F35.[15,16]

One of the more obvious differences between the cryoEM and the crystallographic adenovirus structures is the overall conformation of the penton base.[14–16] In the cryoEM structure the conformation matches the crystal structure of the isolated penton base in complex with an N-terminal fiber peptide,[4] whereas in the X-ray structure of adenovirus the penton base has a more expanded conformation and a larger central pore. In the isolated penton base structure the central pore of the pentamer has a maximum diameter of 28 Å, which is too narrow to accommodate the fiber shaft. In the X-ray structure of the intact Ad5.F35 virion, the penton base pore has an expanded pore diameter of 50 Å and density assigned to the fiber shaft is observed within the pore.[15] In the HAdV5 cryoEM structure, density for a short portion of the fiber shaft is observed on top of the penton base, consistent with the structure of the isolated penton base.[14] It is possible that crystal packing forces helped to induce the altered conformation of the penton base in the Ad5.F35 crystal structure.

The observation of two conformations for the penton base is intriguing. Conformational flexibility of the penton base may play a role in early events in viral cell entry and may be necessary for the programmed disassembly of the virion.[73] It is known that the minor capsid protein VI is membrane lytic and that it is released from the capsid in the endosome during viral cell entry.[74] In the mature adenovirus virion, protein VI is packaged on the inner capsid surface.[16] A conformational change in the capsid, such as dissociation of the penton base, may lead to release of protein VI at the appropriate time during cell entry. A cryoEM study of the adenovirus–integrin interaction led to the hypothesis that strain arising from the symmetry mismatch between four integrin heterodimers and the fivefold penton base might lead to a conformational change in the penton base and promote its release from the capsid.[27]

8. Alternate Assignments for the Four-Helix Coiled Coil

Both the cryoEM and the X-ray structures of adenovirus show four-helix coiled coils at the facet edges (Figure 1(A)).[14–16] Density at this location in the capsid was assigned to a portion of protein IIIa in an early cryoEM analysis of the molecular architecture of adenovirus.[13] This assignment to protein IIIa was based on mass and copy number per capsid. At higher resolution, this density resolved into a four-helix coiled coil, which led to an alternate assignment of this density as the C-terminal domain of protein IX.[12] This new assignment was based on the fact that the C-terminal domain of protein IX is strongly predicted to form a coiled coil and this region was the only observed coiled coil within the icosahedral capsid. This assignment implied that the N-terminal domains of IX form trimers, cementing together hexons within a facet,[30] and the C-terminal domains form four-helix bundles at the facet edges.

The assignment of protein IX to the coiled-coil density at the facet edge seemed to be supported by two moderate resolution cryoEM tagging studies.[75,76] In a tagging study by Marsh et al., an engineered adenovirus with enhanced green fluorescent protein (EGFP) fused to the C terminus of protein IX was examined by cryoEM.[76] The cryoEM structure at 22 Å resolution showed extra density assigned to EGFP at the facet edges hovering above the coiled-coil regions, although these regions were not resolved into separate helices. In a second tagging study by Fabry et al., a 12 residue peptide (called SY12) was engineered at the C terminus of protein IX.[75] A cryoEM structure of the engineered adenovirus at 11 Å resolution showed extra density at both ends of the cylinder of density, representing the coiled-coil region at the facet edge. In addition, anti-SY12 Fab fragments were added and a cryoEM structure of the complex was determined at 22 Å resolution. This structure showed apparent Fab density at both ends of the cylinder of density at the facet edge, indicating that the bundle includes antiparallel helices.

In the atomic resolution cryoEM structure of the intact virion,[14] the four-helix coiled coils were interpreted as the C-terminal domains of protein IX, as assigned earlier by Saban et al.[12] and as indicated by the cryoEM tagging studies.[75,76] The higher resolution cryoEM structure enabled chain tracing within the coiled-coil region with density apparent for several large side chains, including arginines and lysines, which aligned with the atomic model for the C-terminal domain of protein IX.

In the cryoEM-derived atomic model of the four-helix bundle, the helices were linked by a ladder of hydrophobic residues (leucines and valines). Chain tracing indicated that three of the helices were parallel and the fourth was antiparallel. The antiparallel helix was traced as coming from a protein IX N-terminal domain within an adjacent facet. Support for this assignment was provided by the observation that when the cryoEM density was contoured with a low-density threshold, connections were observed between most of the protein IX N-terminal domains and the helices within the coiled coil. The coiled coil is held in place by an interaction with a projecting loop (aa 251–256) on the side of the hexon in position 4 within the capsid.

In the first X-ray structure of the intact virion it was noted that two of the helices in the four-helix bundle appeared to be connected at one end.[15] This density connection between two helices suggested that the four helices might be from a domain of a single protein. This observation, in combination with the lack of clear side-chain density for the helical residues in the coiled coil, led the authors to consider the possibility that this density might be a domain of protein IIIa as originally proposed.[13] In the refined X-ray structure of the virion the assignment of the four-helix bundle to protein IIIa is confirmed with as much certainty as possible given the resolution of the density map (3.8 Å).[16]

9. Protein IIIa Structure

As discussed above, there have been differing assignments for the location of protein IIIa within the capsid. Protein IIIa is the largest cement protein in the capsid (63 kDa) and it is present in 60 copies per virion.[2] It is known to play a role in viral assembly and maturation as temperature-sensitive mutants of protein IIIa are defective for assembly.[77,78] Secondary structure prediction indicates that protein IIIa is highly α-helical with at least 16 predicted helices. Analysis of a cryoEM structure of Ad5.F35 at 6 Å resolution in which α-helices within the capsid were resolved resulted in the assignment of protein IIIa to a cluster of helices below the penton base on the inside of the capsid.[12] A cryoEM labeling study of protein IIIa seemed to support this assignment as it indicated that the N terminus of protein IIIa is located beneath the vertex complex between the penton base and the peripentonal hexons.[79]

The cluster of helices below the penton base was also observed in the atomic resolution cryoEM structure of adenovirus.[14] The backbone fold of a large portion of protein IIIa (aa 7–300) was traced into this density below the vertex. It was reported that side-chain densities were visualized for ~85% of the residues. However, the densities were not distinctive enough to identify individual amino acids and therefore the large side chains were used more as "landmarks" for guiding the building of an atomic model (PDB-ID: 3IYN).

In the refined X-ray structure of adenovirus protein IIIa is assigned to the four-helix bundle on the exterior of the capsid (PDB-ID: 4CWU) (Figure 1(A)).[16] Three segments of protein IIIa are resolved with good certainty. The N-terminal region of protein IIIa (aa 48–102) is observed to extend toward the penton base at the vertex. Another short stretch within the N-terminal region (aa 9–25) is also traced, but the assignment of

this region is less certain. Two segments in the middle of protein IIIa (103–209 and 252–355) are traced within the four-helix bundle at the facet edge. In the atomic model the helical bundle is formed by two long helix–turn–helix motifs with one disordered connection (aa 210–253). In addition, the large C-terminal region of protein IIIa (aa 356–585) is disordered.

Mass spectrometry indicates that the C-terminal 15 residues of HAdV5 protein IIIa (aa 571–585) are cleaved by adenovirus protease,[80] as had been predicted for protein IIIa of HAdV2.[77,81] Reddy and Nemerow surmise that the C-terminal region of protein IIIa remains on the capsid exterior near the icosahedral twofold axis. One remaining puzzle about protein IIIa is how the C-terminal tails are cleaved by the adenovirus protease, which is packaged in the core of the virion.

10. Protein IX Structure

Protein IX is known to help stabilize the virion, as virions lacking protein IX have poor thermostability.[82,83] Recently protein IX has gained prominence as a convenient site of ligand addition for both vector retargeting and fluorescence labeling.[84] The location of the N-terminal domain of protein IX was established by STEM of capsid dissociation fragments called groups-of-nine hexons, or GONs.[30] Four trimeric regions were observed stabilizing the hexon array. Initially these regions were thought to represent the location of the full-length protein IX. However, later it was shown that only the conserved N-terminal domain of protein IX (aa 1–39) is necessary for stabilization of the Ad capsid.[85,86] Volume analysis in a cryoEM study of the Ad5.F35 vector at 9 Å resolution indicated that the locations identified by STEM for protein IX were likely to correspond to only the N-terminal viral interaction domains.[11]

The atomic resolution cryoEM and X-ray crystal structures of adenovirus show density for the N-terminal region of protein IX.[14–16] In the first X-ray structure only coordinates for the Cα backbone atoms were deposited (PDB-ID: 1VSZ).[15] In the cryoEM structure density was visualized for ~85% of the side chains in the N-terminal domain and coordinates were deposited for the majority of the residues in this domain (PDB-ID: 3IYN).[14] Similarly, in the refined X-ray structure coordinates for the N-terminal domain of protein IX were deposited (PDB-ID: 4CWU).[16] However, the N to C direction of the polypeptide backbone is reversed in these two atomic models.

In the refined X-ray structure the best match/confidence scores are obtained for protein IX compared to the scores for the other cement proteins, lending confidence to the X-ray-derived atomic model for protein IX. The protein IX N-terminal regions form triskelion shapes between hexon trimers in a group-of-nine hexons in the middle of each icosahedral facet. In each facet one triskelion sits at the icosahedral threefold axis in the middle of the facet, and three additional triskelions sit at local threefold axes. In the asymmetric unit with just four hexon trimers, one triskelion at a local threefold axis is observed along with one-third of the triskelion at the icosahedral threefold axis (Figure 1(B)). The polypeptide orientation of the

refined X-ray atomic model places the N-termini of protein IX at the distal ends of the triskelion and the middle of the protein IX sequence (~aa 77) at the center of the triskelion.

The protein IX C-terminal domain has a heptad-repeat motif typical of a helix bundle.[87] High-resolution (4–5 Å) cryoEM structures of two bovine adenovirus intermediates showed three-helix coiled coils above the trimeric regions formed by the N-terminal domains of protein IX.[88] No coiled coils are observed in these locations in the human adenovirus structures. In fact, no density at all is observed for the C-terminal domains of protein IX in the refined X-ray structure.[16] The fact that the linker region between the conserved N-terminal region and the predicted C-terminal coiled coil is significantly shorter in bovine adenovirus type 3 (BAdV3) (~24 aa) than in HAdV5 (~42 aa) may explain why a protein IX coiled coil is only observed for BAdV3 and not for human adenoviruses.

A moderate resolution cryoEM structure of the canine adenovirus CAdV2 showed cylinders of density above the protein IX triskelions in the same place as the coiled coils in the BAdV3 structures.[89] As for BAdV3, the linker between the N- and the C-terminal domains of protein IX is significantly smaller in CAdV2 (~15 aa) than in HAdV5 (~42 aa). To help support the assignment of the cylinders to the C-terminal domain of protein IX, Schoehn et al. determined a cryoEM structure of CAdV2 with GFP fused to the C terminus of protein IX.[89] As expected, extra density assigned to GFP was observed above the cylinders. It seems reasonable to conclude that the relatively long linker in HAdV5 protein IX may prevent formation of a rigid coiled-coil bundle extending directly above the N-terminal triskelion region of protein IX.

Given the homology among the N-terminal domains of protein IX among human, bovine, and canine adenovirus, it also seems reasonable to assume that all of these domains have the same fold in the context of intact virions. Assuming that the refined X-ray atomic model is correct,[16] this means that the middle of the protein IX sequence is appropriately placed to have a coiled-coil form above the protein IX triskelion if the linker is short enough. This is apparently the case for both BAdV3 and CAdV2 but not for any of the human adenovirus types that have been studied by cryoEM or X-ray crystallography, including HAdV2, HAdV5, and HAdV12.

11. Core Protein V Structure

One unexpected finding in the refined X-ray atomic model of adenovirus is the positioning of a portion of core protein V on the inner capsid surface (Figure 2(A)).[16] An atomic model was built for 72 residues of protein V (aa 208–219 and 236–295) out of a total of 368 residues. This region of protein V interacts with protein VI below the peripentonal hexons. This positioning is consistent with cross-linking experiments that indicated that proteins V and VI interact within the virion.[90,91] The ordered region of protein V is also observed to interact with the copy of protein VIII that is closest to the vertex. The complex of proteins V, VI, and VIII is observed to stabilize the peripentonal hexons and link them to the adjacent group-of-nine hexons.[16]

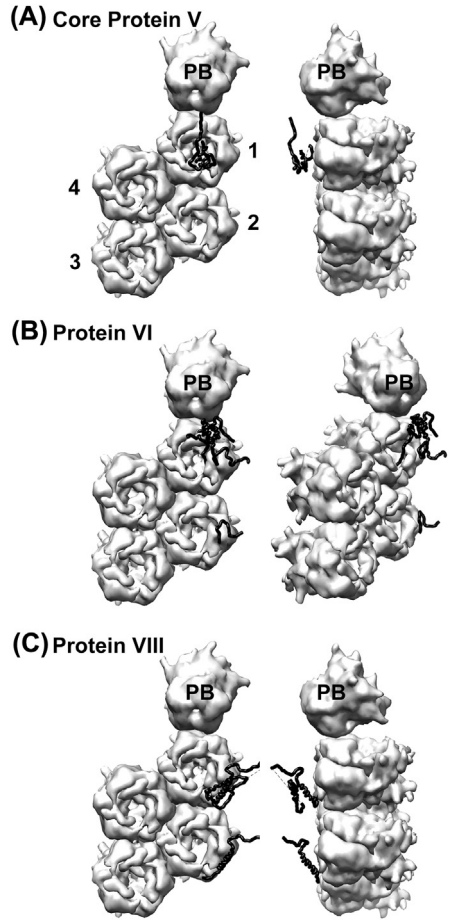

Figure 2 Structure and location of the inner capsid proteins as assigned in the refined adenovirus crystal structure.[16] (A) The enlarged asymmetric unit, with four independent hexon trimers (1–4) and a complete penton base (PB), is shown as in Figure 1 but viewed from the inside of the capsid together with the ordered portion of core protein V (black). Protein V is chain T in PDB-ID: 4CWU. (B) The enlarged asymmetric unit together with the ordered portions of two copies of protein VI (black). The two copies of protein VI in the asymmetric unit are chains U and V in PDB-ID: 4CWU. (C) The enlarged asymmetric unit together with the ordered portions of two copies of protein VIII (black). The two copies of protein VIII in the asymmetric unit are chains X and Y in PDB-ID: 4CWU. Top and side views are shown in panels A and C. Top and a 45° tilted views are shown in panel B. Dashed lines represent disordered regions. This figure was made with UCSF Chimera.[126]

12. Protein VI Structure

Protein VI has multiple functions in the adenovirus lifecycle including regulation of hexon import into the nucleus during adenovirus assembly,[92] disruption of the endosomal membrane during cell entry,[74] and provision of a peptide cofactor for adenovirus

protease.[93,94] During the production of progeny virions in host cells, the viral structural proteins are produced in the cytoplasm while the viral genome is replicated and new viral particles are assembled in the nucleus. Wodrich et al. showed that protein VI shuttles between the nucleus and the cytoplasm and links hexon to the nuclear import machinery via an importin alpha/beta-dependent mechanism.[92] Protein VI contains nuclear import and export signals in a short C-terminal segment, which is proteolytically removed by the adenoviral protease during virus maturation. Wiethoff et al. showed that the N-terminal domain of protein VI has a predicted amphipathic α-helix that is required for membrane lytic activity.[74] Release of protein VI from the virion is thought to occur in the endosome during cell entry. In 1993, two groups showed that an 11-residue peptide cleaved from the C-terminus of the precursor form of protein VI serves as a cofactor for the protease.[93,94]

A direct association between protein VI and hexon has been demonstrated[95,96] and protein VI has also been shown to bind DNA.[97] Therefore a location for protein VI on the inner capsid surface of the virion in the vicinity of the viral genome seems most likely. Also consistent with an internal capsid location is the fact that both the N- and the C-terminal peptide regions of protein VI are cleaved by adenovirus protease. There are ~369 copies of protein VI per virion,[80] which corresponds to ~1.5 copies of protein VI per hexon trimer. Saban et al. first noted density bound within the hexon cavities on the inner capsid surface and tentatively assigned it to protein VI.[11] No coordinates for protein VI were deposited with the atomic resolution cryoEM structure or the first X-ray structure of adenovirus.[14,15]

The refined X-ray structure of adenovirus provided the first atomic model for protein VI (Figure 2(B)).[16] Three regions of protein VI were traced (aa 6–31, 34–79, and 87–157). One copy of protein VI is found within the hexon cavity of each peripentonal hexon. The fold of protein VI appears to be distinct and is predominantly α-helical. The predicted amphipathic α-helix of protein VI[74] does not form an α-helix in the refined X-ray structure. However, it may adopt a helical conformation on interaction with the endosomal membrane.[98] One of the three traced regions (aa 6–31) corresponds to the 33-residue N-terminal propeptide that is cleaved by adenovirus protease. The refined X-ray structure shows that after cleavage the ends of the newly formed fragments are separated by ~24 Å. The majority of the residues in the 33-residue N-terminal propeptide are found with the peripentonal hexon cavity. The propeptide interactions with hexon are consistent with hydrogen–deuterium exchange mass spectrometry results that indicate that the N-terminal propeptide associates with peripentonal hexons.[99] The new structural results are also in agreement with the measured high affinity of the precursor form of protein VI to hexon.[100]

13. Protein VIII Structure

The assignment of protein VIII to two hammer-like regions per asymmetric unit on the inner capsid surface was first made by Fabry et al.[10] These two regions were also observed in the 6 Å resolution cryoEM structure[12] and in the atomic resolution cryoEM and X-ray structures of adenovirus.[14–16] Adenovirus protease cleaves protein

VIII in two places resulting in three fragments. The refined X-ray structure provides coordinates for fragment 1 in both copies of protein VIII within the asymmetric unit (Figure 2(C)).[16] These coordinates mostly agree with the atomic resolution cryoEM coordinates for fragment 1.[14] The refined X-ray structure also includes coordinates for fragment 3 in one copy of protein VIII, although these coordinates differ significantly from the cryoEM coordinates for fragment 3. No density was observed for fragment 2 (aa 112–157) in the refined X-ray structure and it is possible that this fragment is released from the virion after proteolytic processing.

One copy of protein VIII within the asymmetric unit is below the peripentonal hexons. At this position, protein VIII interacts with proteins V and VI and helps to stabilize the interaction between the peripentonal hexons. The second copy of protein VIII is near the icosahedral twofold axis and interacts with protein VI that is bound to the inner side of hexon in position 2 of the asymmetric unit. Both copies of protein VIII are at the edge of a group-of-nine hexons and help to stabilize adjacent facets of hexons.

14. Adenovirus Protease

The adenovirus protease catalyzes the maturational processing of six structural proteins in adenovirus and this step is essential for the production of infectious virus particles.[101,102] These six structural proteins are the precursor forms of proteins IIIa, VI, VII, VIII, mu, and terminal protein (TP).[101,103,104] Three of these are capsid proteins (IIIa, VI, and VIII) and the other three (VII, mu, and TP) are proteins associated with the viral DNA in the core of the virion. Adenovirus protease is also responsible for cleaving the presumed scaffolding protein L1-52K.[105] There are ~50 copies of protease packaged within the core of the virion.[106] Since it plays a critical role in the viral life cycle, adenovirus protease has been proposed as a target for the design and development of antiviral agents to protect against adenovirus infections.[107]

Structures of active[7,108] and inactive[109] forms of adenovirus protease have been determined. The structures confirm the idea proposed earlier that adenovirus protease represents a distinct class of the cysteine proteases.[110] Adenovirus protease was categorized as a cysteine protease on the basis of biochemical and mutagenesis studies.[94,111] Common active site cysteine protease inhibitors are active against adenovirus protease.[106] Originally the sequence of adenovirus protease was unrelated to any other protease sequence in the databases until a weak similarity was found with ubiquitin-like proteinase 1 (Ulp1), which is required for cell-cycle progression in yeast.[112] Recently, two other viral proteases have been added to the adenovirus protease family. These are from vaccinia virus[113] and African swine fever virus.[114] A few other proteins have been found to be homologous to adenovirus protease, including two paralogous gene products in *Chlamydia*,[115] a virulence factor in *Yersinia pestis*, YopJ,[116] and a protease involved in the regulation of chromosome condensation in *Saccharomyces cerevisiae*.[117]

When adenovirus protease is compared to papain, the archetypical cysteine protease, the order of the catalytic Cys and His residues in the primary sequence is different, with His54 followed by Cys122 in adenovirus protease and Cys25 followed by His159