The adenovirus capsid: major progress in minor proteins

Jort Vellinga, Stephanie Van der Heijdt and Rob C. Hoeben

Virus Biology Laboratory, Department of Molecular Cell Biology, Leiden University Medical Centre, Wassenaarseweg 72, 2333 AL Leiden, The Netherlands

Published online ahead of print as DOI 10.1099/vir.0.80877-0.

Human adenoviruses have been the subject of intensive investigation since their discovery in the early 1950s: they have served as model pathogens, as probes for studying cellular processes and, more recently, as efficient gene-delivery vehicles for experimental gene therapy. As a result, a detailed insight into many aspects of adenovirus biology is now available. The capsid proteins and in particular the hexon, penton-base and fibre proteins (the so-called major capsid proteins) have been studied extensively and their structure and function in the virus capsid are now well-defined. On the other hand, the minor proteins in the viral capsid, i.e. proteins IIIa, VI, VIII and IX, have received much less attention. Only the last few years have witnessed a sharp increase in the number of studies on their structure and function. Here, a review of the minor capsid proteins is provided, with a focus on new insights into their position and role in the capsid and the opportunities that they provide for improving human adenovirus-derived gene-delivery vectors.

Introduction

Today, human adenoviruses (HAdV) are among the best-characterized viruses. So far, 51 serotypes have been isolated from humans, which have been classified into six distinct species (previously called subgroups), A–F, according to their sequence similarities and agglutination properties (De Jong et al., 1999; Benkö & Harrach, 2003). The Adenoviridae are a family of non-enveloped viruses of about 90 nm in diameter with a linear, double-stranded DNA genome of 34–48 kb (Russell, 2000; Benkö & Harrach, 2003; Davison et al., 2003). The virus capsid contains at least nine proteins, of which hexon (protein II), penton base (protein III) and fibre (protein IV) are termed major capsid proteins, and proteins IIIa, VI, VIII and IX minor capsid proteins (Fig. 1). Historically, these proteins were numbered (II–IX) in order of their increasing electrophoretic mobilities of purified HAdV-2 particles on SDS–polyacrylamide gels. The other four proteins, protein V, protein VII, µ and the terminal protein TP, are packaged with the genomic DNA in the viral core (Everitt et al., 1973; Chatterjee et al., 1985; van Oostrum & Burnett, 1985). The development of efficient gene-delivery vectors derived from HAdV has led to a renewed interest in their capsid structure and biology. Whereas the HAdV-C serotypes 5 (HAdV-5) and 2 (HAdV-2) initially dominated the research, the focus has been expanded and now also involves other species (i.e. HAdV-B, D and E). This was triggered by the notion that the diversity of HAdV could be exploited to expand the applicability of HAdV vectors. Examples of such developments are the exploration of HAdV-B-derived vectors to overcome the scarcity of the coxsackievirus–adenovirus receptor (Stevenson et al., 1995; Gall et al., 1996; Havenga et al., 2002; Uil et al., 2003), which is used by HAdV-C-derived vectors, and the selection of HAdV serotypes with a low prevalence of neutralizing immunity (Vogels et al., 2003; Barouch et al., 2004; Holterman et al., 2004).

The main strategy for manipulating the vector’s tropism has been the insertion of receptor-binding ligands into the major capsid proteins. Such ligands were inserted at exposed positions in the viral capsid, to recruit alternative receptors on target cells. The fibre, penton base and hexon have all been exploited for virus retargeting and many studies have demonstrated the feasibility of the approach (Crompton et al., 1994; Huvent et al., 1998; Einfeld et al., 1999; Vigne et al., 1999; Krasnykh et al., 2001; Mizuguchi et al., 2001; Barnett et al., 2002; Belousova et al., 2002; Wu et al., 2005). These studies have been facilitated by the availability of the crystal structure of these proteins. The preferred positions for such insertions of targeting ligands are in the HI loop of the fibre knob, in the hypervariable region 5 of hexon loop L1 and in the RGD-motif loop of penton base. This has been covered in an excellent review by Mizuguchi & Hayakawa (2004). However, structural constraints of these surface-exposed loops limit their tolerance for larger heterologous peptides. In addition, oligopeptide ligands have been added to the C terminus of the fibre (Michael et al., 1995; Wickham et al., 1996). Also, stringent size limitations were imposed by the structure of the fibre protein on the ligands...
incorporated and thus limited the range of potential ligands to short peptides. In a bolder approach, an artificial trimerization domain coupled to globular receptor-binding ligands has replaced the entire knob domain of the fibre. Although the feasibility of this approach has been demonstrated by several groups, the resulting vectors are difficult to produce at high titres and the artificial fibres are frequently inserted inefficiently into the capsid (Krasnykh et al., 1996; van Beusechem et al., 2000; Henning et al., 2002; Magnusson et al., 2002).

In this respect, the minor capsid proteins have received less attention. Lately, however, there has been a renewed interest in their function and their use in HAdV-based vector design. Modification of the minor capsid proteins has recently been used successfully for HAdV and bovine AdV targeting (Curiel, 2002; Dmitriev et al., 2002; Vellinga et al., 2004). Here, we will review the minor capsid proteins, with a focus on their role and position in the capsid and the opportunities that they provide for modifying HAdV-derived gene-delivery vectors.

**Adenoviral capsid composition**

The HAdV particle is composed of more than 2700 polypeptide molecules, with a total mass of approximately $150 \times 10^6$ Da. Determining their precise three-dimensional structure was a challenging task. Nevertheless, several cryo-electron microscopic studies using image-reconstruction techniques have provided a detailed three-dimensional model of the adenovirus capsid (Stewart et al., 1993; Rux & Burnett, 2000; Rux et al., 2003; San Martin & Burnett, 2003). These data confirmed and extended results of earlier studies in which enzymic iodination, immunoprecipitation, chemical cross-linking techniques and Ar$^+$ plasma etching were used to establish the specific order of the polypeptides and their positions in the capsid (Everitt & Philipson, 1974; Everitt et al., 1975; Newcomb & Brown, 1988) (Fig. 1).
Hexon is the largest and most abundant of the structural proteins in the adenovirus capsid. The 720 hexon monomers present in each virion form 240 hexon homotrimeric, which in turn form 20 capsid facets, each consisting of 12 hexon homotrimers (van Oostrum & Burnett, 1985; Stewart et al., 1993). The other two major capsid proteins, fibre and penton base, form the penton complexes (three subunits of protein IV and five subunits of protein III) at each virion vertex. The position of the minor capsid protein IIIa has been mapped to the icosahedral edges and hexamers of protein VI are positioned underneath the penton base (Stewart et al., 1991, 1993) (Fig. 1a, c). Here, protein VI interacts with a core protein, protein V, which is located internally near the penton base in the vertex region of the virion. Cryo-electron microscopic image analyses, combined with groups-of-nine hexons (GON) modelling of the hexon crystal structure, determined the position for the smallest minor capsid protein, protein IX (Fig. 1b, c): four trimeric densities were found near the outer surface of the capsid, in the central cavities between the hexon capsomers that form the GONs (Furcinitti et al., 1989; Stewart et al., 1991). There is no clear evidence of the minor capsid protein VIII on the outer surface of the capsid; therefore, this protein is presumed to be located internally in the capsid (Stewart et al., 1991, 1993).

**Minor capsid proteins**

**Protein IIIa**

The precursor of protein IIIa has a mass of 67 kDa and is cleaved at the N terminus during maturation of the virion to generate the 63-5 kDa protein IIIa. At this time, it can be detected in the excess pool of virus material. The presence of the protein in the capsid is biologically significant, as temperature-sensitive (ts) mutants that affect protein IIIa (viz. H2ts4, H2ts112, H5ts58 and H2ts101) are defective in the assembly of virions and accumulate empty and protein IIIa-deficient particles at the non-permissive temperature (D’Halluin et al., 1982; Chroboczek et al., 1986). The empty particles formed by H2ts101 and H2ts112 at the non-permissive temperature are assembly intermediates, and mature to form complete particles when incubated at the permissive temperature. The defect may be caused by defective cleavage, which blocks the conversion of the precursor protein to mature protein IIIa at the non-permissive temperature (Boudin et al., 1980; D’Halluin et al., 1982). Imaging techniques have shown that protein IIIa is present as an elongated, monomeric polypeptide extending from the outer surface through to the inner space bounded by the capsid (Fig. 1a). Evidence that protein IIIa domains interact with protein VII comes from experiments showing that proteins IIIa and VII can be co-immunoprecipitated (Boudin et al., 1980; Stewart et al., 1993). The main density assigned to protein IIIa is located near the outer surface of the capsid, where protein IIIa may stably assemble the facets (Stewart et al., 1993; San Martin & Burnett, 2003) (Fig. 1). Protein IIIa is located at the surface of the particle and its modification may be a means of modifying the tropism.

As a first step for evaluating this approach, a six-His tag was incorporated into the N terminus of protein IIIa. The presence of the inserted tag in the gene encoding protein IIIa could be detected by PCR in DNA extracted from viral plaques, but no data were provided on the presence and accessibility of the tags in intact viral particles (Curiel, 2003).

At least part of the protein IIIa molecules are phosphorylated in the capsid (Tsuzuki & Luftig, 1983, 1985; San Martin & Burnett, 2003). The physiological impact of this phosphorylation is not known.

An additional function for protein IIIa has recently been proposed. The adenovirus major promoter is also active early during infection and, at this stage, only one mRNA is produced: the L1 52,55K mRNA. However, this transcript is alternatively spliced. The last intron is spliced using a common 5’ splice site and two competing 3’ splice sites. This generates two predominant cytoplasmic mRNAs: the 52,55K (with the proximal 3’ splice site) mRNA and the IIIa (with the distal 3’ splice site) mRNA, respectively. Generation of the IIIa mRNA occurs only after the onset of DNA replication. Recently, it was observed that illicit expression of the protein IIIa gene reduced the formation of the L1 52,55K mRNA and strongly increased formation of IIIa mRNA (Molin et al., 2002). These data suggest that, late in infection, protein IIIa may enhance its own synthesis by an autostimulatory mechanism at the expense of the L1 52,55K mRNA by affecting the splicing machinery. The mechanism is as yet unclear.

**Protein VI**

The mature version of protein VI is 22 kDa and is generated by cleavage from a larger precursor (pVI). The maximum rate of synthesis in the cytoplasmic fraction is at 14–15 h post-infection (Everitt & Philipson, 1974). Protein VI is positioned in the interior of the capsid, presumably adjacent to the hexons (Fig. 1a). Here, it may connect the bases of two neighbouring peripental hexons (Everitt et al., 1975; Stewart et al., 1991, 1993; Greber et al., 1993) (Fig. 1c).

It has been suggested that 114 residues of the central part of the protein VI are ‘disordered’. Crystallographic studies have shown that similar domains in other viral capsid proteins interact with the viral core. Molecular reconstruction revealed protein VI densities that connect to the core (Stewart et al., 1991). The similarity of the basic regions of protein VI to those of other viral proteins suggests that the basic region of protein VI may interact directly with viral DNA (San Martin & Burnett, 2003; Stewart et al., 1993).

During infection, protein VI may aid the virion particle to escape from the endosome. Here, protein VI induces a pH-independent disruption of the membrane (Wiethoff et al., 2005). Further proteolysis of protein VI occurs after endosome lysis and entry into the cytoplasm. The proteolysis is catalysed by the virally encoded DNA-associated cysteine protease. This protein, the so-called ‘adenaine’, can use several viral proteins as a substrate (i.e. protein VI and...
the precursor proteins pVII, pVIII, pIIa and pTP (Greber et al., 1993; Webster et al., 1993). Intriguingly, the carboxy-terminal 11 aa peptide, which is cleaved from pVI and is termed pVI-CT, forms a disulphide-linked heterodimer with adenaine and activates the protease activity (Matthews et al., 1993; Webster et al., 1993).

An important function of protein VI is to facilitate nuclear import of hexon proteins. In cells infected with H2ts147, a mutant defective in protein VI, no nuclear import of hexon is observed at the non-permissive temperature (Matthews & Russell, 1995; Wodrich et al., 2003). The precursor of protein VI (pVI) acts as a shuttle, facilitating the hexon transfer into the nucleus by recruiting the importin α/β-dependent system. Protein pVI contains two nuclear-localization sequences (NLS) and two nuclear-export sequences (NES). NLS2 and NES2 are removed by proteolysis during maturation. Studies with deletion mutants showed that NLS2 is important for nuclear localization. When a mutant containing only NLS2 was injected into a nucleus, it remained there. In contrast, if a mutant containing both NLS2 and NES2 was injected into a nucleus, nuclear–cytoplasmic shuttling was observed. It has been suggested that the NES is masked by hexon and thus is inaccessible (Wodrich et al., 2003), preventing nuclear export of pVI bound to hexon. This is supported by the fact that a AC-terminal mutant of pVI, just like the mature, processed form, fails to facilitate nuclear import of the hexons (Matthews & Russell, 1995; Wodrich et al., 2003).

Protein VIII

The least-studied of the minor capsid proteins is the 15·3 kDa protein VIII. It is located at the inner surface of the triangular facets as dimers and interacts with hexons of adjacent facets (Fig. 1a). Like proteins VI and IIIa, it is synthesized as a larger precursor protein that is processed proteolytically by adenaine (Stewart et al., 1993). Whereas the precursor of protein VIII is present in empty capsids, it is undetectable in complete particles. This feature has been used to assay for the presence of empty particle in HAdV-5 vector preparations (Vellekamp et al., 2001). It has a high content of serine (17–20 %), proline (8 %) and basic residues arginine and lysine (11 %), which suggests that it may be largely disordered (Stewart et al., 1993). Analyses using mutant viruses suggest that protein VIII plays a role in the virion’s structural stability (Liu et al., 1985).

Protein IX

The 14·3 kDa protein IX is the smallest of the minor capsid proteins. In contrast to the other minor capsid proteins, protein IX is unique to the mastadenoviruses and is absent in the other adenoviruses genera. Twelve molecules of protein IX are located at each of the 20 facets of the icosahedral capsid. With nine hexon capsomers, protein IX forms the stable assemblies that are termed GONs (Fig. 1b, c). These GONs form the central part of each of the facets of the icosahedral capsid (Everitt & Philipson, 1974; Everitt et al., 1975). The protein IX molecules have been positioned in the cavities between the hexon tops, where they form continuous trimeric densities (Furcinitti et al., 1989; Stewart et al., 1991, 1993) (Fig. 1a). This led to the conclusion that protein IX resides in the capsid as trimers. The conserved leucine-zipper domain in the carboxyl-terminal part of the protein allows protein IX to self-associate (Rosa-Calatrava et al., 2001; Vellinga et al., 2005). This region may form a trimeric coiled-coil region, stabilizing the trimers. Although mutant HAdV-5 viruses lacking protein IX can be propagated with approximately the same efficiency and titres as wild-type (wt) HAdV-5, they do not form GONs and the virions are more heat-sensitive than wt HAdV particles (Boulanger et al., 1979; Colby & Shenk, 1981). Hence, protein IX has been described to act as capsid cement (Furcinitti et al., 1989). The current model of the adenovirus capsid is firmly established. Although the location of protein IX in a central position in the GONs is conceptually very elegant and widely accepted, new data obtained with sophisticated imaging techniques suggest an alternative location of protein IX within the capsid (Campos et al., 2004b). Image analyses of virions with extended protein IX molecules position protein IX near the peripental hexons. In the classical model, this position is occupied by protein IIIa. This is an intriguing observation, but it is difficult to envisage how protein IX in this position can stabilize the GONs.

The N terminus of protein IX is strongly conserved between serotypes. This region is required for insertion of protein IX into the capsid and for the thermostable phenotype of the HAdV-5 particles. In contrast, deletions that affect the other two conserved regions of protein IX, i.e. the central alanine-rich region and the C-terminal coiled-coil domain, impair neither protein IX incorporation into the capsid nor the thermostability of the resulting virions (Rosa-Calatrava et al., 2001; Vellinga et al., 2005). This demonstrates that protein IX trimerization is dispensable for its incorporation into the capsid and for capsid stabilization (Vellinga et al., 2005).

In addition, protein IX affects the DNA-packaging capacity of HAdV. Whereas virions harbouring protein IX in their capsids can accommodate 1.5–2·0 kb DNA in excess of the normal genome length (i.e. 105 % of the normal length), virions without protein IX have a DNA-carrying capacity that is approximately 2 kb less than the normal length. Even the wt HAdV genomes are packaged inefficiently (Ghosh-Choudhury et al., 1987). This feature has recently been exploited in a new system for generation of helper-dependent adenovirus vectors. The combination of protein IX-deleted helper viruses with a Cre-mediated excision in the viral-packaging signal resulted in a 1000-fold reduction of helper-virus contamination (Sargent et al., 2004b). Moreover, the authors provide evidence that capsids lacking protein IX can accommodate at least 37·3 kb viral DNA, but that the resulting virions are not infective in plaque assays. This is an intriguing observation, although the mechanism by which infectivity is lost remains elusive.
The leucine-zipper region of protein IX is involved in the formation of so-called protein IX (inclusion) bodies during infection (Rosa-Calatrava et al., 2001; Souquere-Besse et al., 2002). After infection, the ultrastructure of the nucleus changes. Nuclear inclusion bodies containing protein IX are formed. The cellular promyelocytic leukaemia (PML) protein is relocated to these inclusions late in infection. This PML protein is involved in regulating the cellular antiviral response, which could explain why it would be beneficial for the virus to contain this protein in inclusion bodies (Rosa-Calatrava et al., 2003). There is evidence that the leucine-zipper domain is indispensable for the formation of inclusion bodies, because mutations in this domain abolish inclusion bodies. These data suggest that the leucine-zipper domain is involved in generating the protein IX bodies. It will be very interesting to identify the cellular partners associating with protein IX. This may provide insight into the functional significance of protein IX’s capacity to self-associate via the leucine-zipper domain.

In addition, protein IX has been shown to affect transcriptional activity of several promoters.Transient-expression assays demonstrated that protein IX stimulates the AdV E1A, E4 and the major late promoter activity. For this stimulating activity, the integrity of the leucine-zipper domain and the central alanine stretch is essential (Lutz et al., 1997; Rosa-Calatrava et al., 2001). More recent studies, however, show that protein IX has only a modest effect on the activity of the E1A and E4 promoters (Sargent et al., 2004a). Moreover, in 293 cells, the transcription-stimulatory role of protein IX is not essential for adenovirus replication (Sargent et al., 2004a). These apparently contradictory results may merely reflect differences in the cell systems used for evaluating the effects of protein IX on transcriptional activity. Studies of infection of diploid cells in primary cultures or in suitable animal models might resolve this issue in the future. It would, therefore, be interesting to study the behaviour of wt and protein IX-mutant HAdV in cotton rats. Cotton rats are at least semipermissive for HAdV-5 infection and have been used as an animal model for studying adenovirus pathogenesis (Pacini et al., 1984), and therefore provide an in vivo system to study HAdV biology in an immunocompetent host.

Immunooaccessibility studies with antibodies directed against the N-terminal or the C-terminal part of HAdV-3 protein IX suggest that the N terminus lies hidden between the hexon capsomers. In contrast, the C-terminal domain is accessible to immunoglobulins, suggesting that it is pointing towards the outer surface (Akalu et al., 1999). This observation triggered several studies to evaluate the use of protein IX as an anchor for heterologous ligands. If these ligands could recruit new receptors, they can be used to expand the viral tropism. To this end, FLAG epitopes and poly-lysine tracts were fused with the C terminus of protein IX (Dmitriev et al., 2002). Poly-lysine residues bind heparan sulphates at the cell surface. Anti-FLAG antibodies could bind to the tags on intact particles, demonstrating that the FLAG epitopes linked to protein IX were accessible in the capsid. In addition the poly-lysine tract could bind to heparin-coated beads (Dmitriev et al., 2002). In another study, x-helical spacers of various lengths were placed between the C terminus of protein IX and a ligand. The accessibility of these ligands correlated with the spacer length, demonstrating that the spacers could be used to lift the ligands and expose them at the capsid surface (Vellinga et al., 2004). A protein IX–green fluorescent protein (GFP) fusion with a FLAG epitope was generated to prove that large proteins fused to protein IX could be integrated in the capsid. The amount of protein IX–GFP incorporated into virions was similar to that of protein IX–FLAG (Le et al., 2004; Meulenbroek et al., 2004). Also, the particle : p.f.u. ratios of batches of these viruses were similar, demonstrating that protein IX–GFP can be incorporated into the capsid without affecting the stability of capsid. The GFP-labelled virions have been used to monitor and track the virions in vitro and in vivo (Le et al., 2004; Meulenbroek et al., 2004). Furthermore, coupling of a biotin-acceptor peptide to the C terminus of protein IX has been shown to facilitate binding of biotinylated ligands on the adenovirus capsid surface (Campos et al., 2004a). This technique provides a flexible and potentially robust technique for capsid modification.

The accessibility of the ligands linked to the C terminus of protein IX, together with the demonstration that trimer formation of protein IX is required neither for its capsid incorporation nor for generating thermostable viruses, make pIX a promising target for capsid modification. It may not be essential to retain its capacity to form trimers when modifying protein IX for vector retargeting. This is in contrast to the fibre, for which trimerization is essential for its association with penton base.

A reappraisal of the minor capsid proteins

Recently, protein IX has become a central focus of the adenovirology field. It is reasonable to anticipate that the other minor capsid proteins also will reemerge. Although proteins VI and VIII are located at the inner surface of the capsid and therefore cannot be used to modify viral tropism, their precise function in the genesis of the viral capsid remains an interesting issue, for which many details have to be filled in. It is to be expected that, in the near future, new structural analysis will provide precise details of their structure, position and role in the capsid. The renewed interest in the structure and function of the minor capsid proteins is fully justified. Recent studies have already provided new and challenging data and form a firm foundation for further study. It is also good to realize that, with the proteins of the adenovirus capsid, it is function that matters, rather than size.

Acknowledgements

We thank Drs Hans van Ormondt and David Baker for stimulating discussions and for critical reading of the manuscript.
References


