Virology 385 (2009) 434–443

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Adeno-associated virus-2 and its primary cellular receptor—Cryo-EM structure of a heparin complex

Jason O'Donnell [a], Kenneth A. Taylor [a], Michael S. Chapman [b,*]

[a] Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306-4380
[b] Department of Biochemistry and Molecular Biology, School of Medicine, Mail code L224, Oregon Health and Science University, 3181 S.W. Sam Jackson Park Road, Portland, OR 97239-3098

ARTICLE INFO

Article history:
Received 6 October 2008
Returned to author for revision
10 November 2008
Accepted 18 November 2008
Available online 13 January 2009

Keywords:
Heparan sulfate
Heparin
Gene therapy
Electron microscopy
Parvovirus

ABSTRACT

Adeno-associated virus serotype 2 (AAV-2) is a leading candidate vector for gene therapy. Cell entry starts with attachment to a primary receptor, Heparan Sulfate Proteoglycan (HSPG) before binding to a co-receptor. Here, cryo-electron microscopy provides direct visualization of the virus–HSPG interactions. Single particle analysis was performed on AAV-2 complexed with a 17 kDa heparin fragment at 8.3 Å resolution. Heparin density covers the shoulder of spikes surrounding viral 3-fold symmetry axes. Previously implicated, positively charged residues $R_{448/585}$, $R_{451/588}$ and $R_{350/487}$ from another subunit cluster at the center of the heparin footprint. The footprint is much more extensive than apparent through mutagenesis, including $R_{347/484}$, $K_{395/532}$ and $K_{390/527}$ that are more conserved, but whose roles have been controversial. It also includes much of a region proposed as a co-receptor site, because prior studies had not revealed heparin interactions. Heparin density bridges over the viral 3-fold axes, indicating multi-valent attachment to symmetry-related binding sites.

Adeno-associated virus serotype 2 (AAV-2) was first discovered in tissues of children infected with adenovirus, from which its name is derived (Laughlin et al., 1983; Salo and Mayor, 1977; Salo and Mayor, 1979). Since its discovery, AAV-2 has emerged as a leading candidate vector for gene therapy. It was considered well-suited for development as a vector because the wild-type virus is non-pathogenic, elicits a low immune response, integrates itself site-specifically into chromosome 19, infects both dividing and non-dividing cells and has wide cell tropism (Berns and Giraud, 1995; Bueler, 1999; Chirmule et al., 1999; Podsakoff et al., 1994; Trempe, 1996; Wu et al., 1998, 2006a, 2006b). Recombinantly engineered AAV vectors (rAAV) differ from wild-type in that transduction is primarily episomal, and that the vector used is non-replicative and therefore not dependent on the presence of helper adenovirus (Carter, 1990; Penaud-Budloo et al., 2008). Experimental vectors have been designed to combat a variety of diseases such as AIDS, high blood pressure, cystic fibrosis, and Parkinson's disease (Chaterjee et al.,1992; Chen et al., 1996; Flotte and Carter, 1998; Harster et al., 1999; Inouye et al., 1997; Mandel et al., 1997; Mochizuki et al., 2001; Phillips, 1997). However, AAV's promiscuous cell infection may also be considered a disadvantage and an obstacle towards the achievement of tailored cell-specific vectors. Structural information regarding host–cell interactions could accelerate this development.

AAV-2 is a dependovirus within the parvovirus family and is the type species. It has an ~4.7 kb ssDNA genome encapsulated by a $T=1$ icosahedral capsid built from three proteins, VP1, VP2, and VP3 in a respective ratio of 1:1:10 (Caspar and Klug, 1962; Xie et al., 2002). Common to all subunits is the 533 residue component, corresponding to VP3, which was seen in the crystal structure of AAV-2, determined to 3 Å (Xie et al., 2002). It was the infectious particles of AAV-2 that were crystallized, containing the usual complement of VP1 and VP2, but the unique parts of these minor capsid proteins were not observed in the structure which was averaged according to icosahedral symmetry (Xie et al., 2004, 2003). Relative to VP3, VP2 and VP1 are extended by 65 and 202 residues respectively at the N-terminus. The unique part of VP1, "VP1u" carries a phospholipase domain that is important in transport of the viral genome from endosomes and/or into the nucleus (Girod et al., 2002). Two conventions for residue numbering are well-established in the literature, one counting from the N-terminus of VP2, the other from VP1 and differing by 137 (for AAV-2). Here, to minimize ambiguity, both will be provided in the form VP2/VP1.

The common part of the VP1/2/3 subunit has a jelly-roll β-barrel motif with long-loop insertions between the core β-barrel strands. These intra-strand loops compose 60% of the structure and most of the distinctive surface topology. Loops from neighboring subunits interdigitate to form 3-fold related peaks as well as the putative receptor binding site found along the shoulder of each peak. These loops have a lower sequence identity (<15%) amongst different parvovirus genera

Abbreviations: AAV, Adeno-associated virus; CTF, contrast transfer function; EM, Electron microscopy; e.u., error units (calculated from the map variance between two half data sets, see text); FMDV, Foot and mouth disease virus; HS, heparan sulfate; HSPG, heparan sulfate proteoglycan.

* Corresponding author. Fax: +1 503 494 8393.
E-mail address: chapmami@ohsu.edu (M.S. Chapman).

0042-6822/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2008.11.037

and probably determine to a large extent tissue tropism (Chapman and Rossmann, 1993). AAV-2's infection pathway is common to many viruses.

On entering many types of cells, AAV-2 first binds a low affinity attachment (primary) receptor, heparan sulfate proteoglycan (HSPG), to localize on the cell surface and then requires the presence of a co-receptor for endocytosis and infection (Summerford and Samulski, 1998). Four co-receptors have been proposed to date for AAV-2 and its close relatives: fibroblast growth factor receptor (FGFR), hepatocyte growth factor receptor (HGFR), the 36/67 kDa Laminin receptor and integrins $\alpha_V\beta_1$ and $\alpha_V\beta_5$, though the latter has been questioned (Akache et al., 2006; Asokan et al., 2006; Blackburn et al., 2006; Kashiwakura et al., 2005; Qing et al., 1999; Qiu and Brown, 1999; Smith et al., 2004; Summerford et al., 1999). There are alternatives to AAV-2's HPSG-mediated cell entry, as revealed through tissue-dependent effects of mutation at heparin-binding motifs, poor correlation of infectivity with HSPG expression, and transduction by AAV vectors of low HSPG cells through CD9-mediated entry (Kern et al., 2003; Kurzeder et al., 2007; Qiu et al., 2000). Other serotypes, AAV-1, -4, -5, and -6, bind to sialic acid with different specificities for its various forms (Kaludov et al., 2001; Walters et al., 2001; Wu et al., 2006c). Thus, while cell entry is more complicated than once presumed, the consensus is that HSPG-binding plays an important, often dominant, role for at least AAV-2 and AAV-3 (Kurzeder et al., 2007).

HSPG is found on the surface of many cells. It is comprised of an integral membrane protein with an attached heparan sulfate (HS). HS, like heparin, is a linear polysaccharide glucuronic/iduronic acid and glucosamine residues. The N- and O-linked sulfates of such heparinoids impart a negative charge which is implicated in their biological interactions with growth factors, chemokines, proteases, and pathogens (Conrad, 1998). HS differs from heparin in a higher diversity of disaccharides, lower levels of sulfonation and greater ubiquity, but heparin which is more specific to mast cells is a close analog that is prepared for pharmaceutical use and is commonly used as a structural and functional analog (Conrad, 1998). Examples of viruses that use HS for cellular attachment include Foot and Mouth disease virus (FMDV), Herpes Simples Virus (HSV-1), human immunodeficiency virus (HIV-1), respiratory syncytial virus, Dengue virus, and several alphaviruses (Capila and Linhardt, 2002; Chen et al., 1997; Conrad, 1998; Jackson et al., 1996; Krusat and Streckert, 1997; Patel et al., 1993; WuDunn and Spear, 1989; Zhang et al., 2005). Atomic structures of Heparin–protein complexes have shown that interactions are often dominated by multiple ionic pairs between arginines and heparin's sulfate groups while polar interactions with sugar hydroxyls are also common (Carfi et al., 2001; Dementiev et al., 2004; DiGabriele et al., 1998; Faham et al., 1996; Lietha et al., 2001; Mulloy and Linhardt, 2001). Beyond this, there is little conservation of a binding motif between different HS-binding proteins so recognition of binding sequences is difficult (Hileman et al., 1998). In the report of the crystal structure of FMDV subtype 01, complexed with heparin, it was suggested that a single heparin fragment could bridge between multiple receptor sites on the symmetrical virus capsid (Fry et al., 1999). It was proposed that weak binding interactions of individual sites could be combined to achieve strong binding in sum.

The crystal structure of AAV-2, and calculation of the surface electrostatic potential suggested that the HS binding site might be a patch of strong positive charge on the shoulder of surface protrusions where 4 arginines and a lysine come together (Xie et al., 2002). Subsequently, three independent mutagenesis studies tested this postulate, and all concluded that basic residues on the capsid surface were crucial for HSPG binding and cell infection (Kern et al., 2003; Lochrie et al., 2006; Opie et al., 2003). There is consensus that a cluster of positively charged residues on the side of each 3-fold peak, specifically $R_{448/585}$, $R_{451/588}$ and $R_{350/487}$ are involved in heparin binding. Unresolved issues include the lack of conservation of, $R_{451/588}$ and $R_{350/487}$ among heparin-binding serotypes, and differing experimental characterizations of mutations at more conserved sites, including $R_{347/484}$, $R_{350/487}$ and $K_{395/532}$.

Reported here are cryo-EM reconstructions of AAV-2 with and without a 17 kDa receptor analog fragment at resolutions of 8.3 Å and 7.8 Å respectively. Difference map analysis shows the strongest peak as an elongated ellipsoidal shape laying tangential to capsid surface directly over the basic residues $R_{448/585}$ and $R_{451/588}$ and extending toward $R_{347/484}$, $R_{350/487}$ and $K_{395/532}$, in fact enclosing a footprint much larger than previously appreciated. At lower contour levels, difference density extends between 3-fold and also 2-fold related binding sites, indicating a single polysaccharide receptor bridges between multiple sites which together presumably achieve tighter attachment.

Results and discussion

Quality of the electron microscopic reconstructions

AAV-2 complexed with a 17 kDa receptor analog fragment, together with a native control, has been visualized through cryo-EM reconstruction (Figs. 1A, C) at resolutions of 8.3 Å and 7.8 Å respectively (Fig. 2) that are among the highest that have been obtained for any parvovirus (Kaufmann et al., 2008). The reconstruction of the native virus was built from 7491 particle images.

The agreement between the EM reconstruction of the native virus and the crystallographic structure is excellent, even though perfect agreement should not be expected at 8 Å resolution. All surface features characteristic of AAV-2 are apparent, including conspicuous 3-fold related peaks, a 2-fold dimple and an empty 5-fold channel. The overall consistency is remarkably good, and improved relative to the earlier EM visualizations of AAV-2 particles at *ca.* 10 Å resolution (Kronenberg et al., 2005). Several hairpin loops, comprised of just two antiparallel backbone chains are clearly visible on the surface, such as the EF loop between β-strands E and F, five symmetry-equivalent copies of which form distinctive outer surface features surrounding the 5-fold axis (Fig. 2).

Interpretation was restricted to map features exceeding 3 or 4 error units. This threshold appears to be robust from several perspectives: (1) the mock difference map (odd–even images) gives highest peaks in general positions of ±3 e.u. (noting that symmetry axes can have higher noise). (2) The native EM map, contoured at 3.8 e.u. gives good correspondence to the crystallographic structure — at 4.4 e.u. some parts are outside the density, while at 1.8 e.u. all is enclosed within a possibly over-generous envelope. (3) If a single 17 kDa fragment were bound at each of 60 symmetrical sites, the appropriate volume would be enclosed by an isocontour at 2.5 e.u., 4.5 e.u. if each heparin bridges between two binding sites. Thus, three different criteria suggest that the transition from noise to real features occurs between +2.5 and +4.5 e.u.

The reconstruction of the HS-complex was built from 5950 individual virion images (Figs. 1B, D). Overall, there is a good agreement with the native reconstruction. Extra density in the map of the complex covers the entire 3-fold area. The differences between heparin-complex and native are most apparent in the difference map which is interpreted below, but all of the significant features could also be seen in the map of the complex (Fig. 1).

Visualization of the bound heparin receptor analog

In the heparin-native difference map, the strongest density at +6.9 e.u. is located just off the viral surface near a concentration of 6 positively charged residues. The closest is $R_{448/585}$, about 5 Å away from the highest density (Table 1 and Figs. 3 and 4). Strong density > +6 e.u. wraps around neighboring side chains, $R_{448/585}$ and $R_{451/588}$ (Table 1), which are on the inner shoulder of each peak surrounding the 3-fold axes. Density at +5.2 e.u extends from each of the shoulders, connecting over and around the 3-fold axes, bridging the 28 Å wide valley between 3-fold-related peak shoulders. If the contour level is