Virology 378 (2008) 169–176

Contents lists available at ScienceDirect

Virology

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

Structure of the C-terminal head domain of the fowl adenovirus type 1 short fibre

Majida El Bakkouri [a], Elena Seiradake [b,1], Stephen Cusack [b], Rob W.H. Ruigrok [a,*], Guy Schoehn [a,c,*]

[a] Unit for Virus Host Cell Interactions, UMR5233 UJF-EMBL-CNRS, CIBB, 6 rue Jules Horowitz, BP181, 38042 Grenoble Cedex 9, France
[b] EMBL Grenoble Outstation, 6 rue Jules Horowitz, BP181, 38042 Grenoble Cedex 9, France
[c] Institut de Biologie Structurale Jean–Pierre Ebel, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France

ARTICLE INFO

Article history:
Received 4 April 2008
Returned to author for revision
1 May 2008
Accepted 14 May 2008
Available online 17 June 2008

Keywords:
Structure
X-ray crystallography
CELO short fibre head domain
adenovirus

ABSTRACT

There are more than 100 known adenovirus serotypes, including 50 human serotypes. They can infect all 5 major vertebrate classes but only Aviadenovirus infecting birds and Mastadenovirus infecting mammals have been well studied. CELO (chicken embryo lethal orphan) adenovirus is responsible for mild respiratory pathologies in birds. Most studies on CELO virus have focussed on its genome sequence and organisation whereas the structural work on CELO proteins has only recently started. Contrary to most adenoviruses, the vertices of CELO virus reveal pentons with two fibres of different lengths. The distal parts (or head) of those fibres are involved in cellular receptor binding. Here we have determined the atomic structure of the short-fibre head of CELO (amino acids 201–410) at 2.0 Å resolution. Despite low sequence identity, this structure is conserved compared to the other adenovirus fibre heads. We have used the existing CELO long-fibre head structure and the one we show here for a structure-based alignment of 11 known adenovirus fibre heads which was subsequently used for the construction of an evolutionary tree. Both the fibre head sequence and structural alignments suggest that enteric human group F adenovirus 41 (short fibre) is closer to the CELO fibre heads than the canine CAdV-2 fibre head, that lies closer to the human virus fibre heads.

Introduction

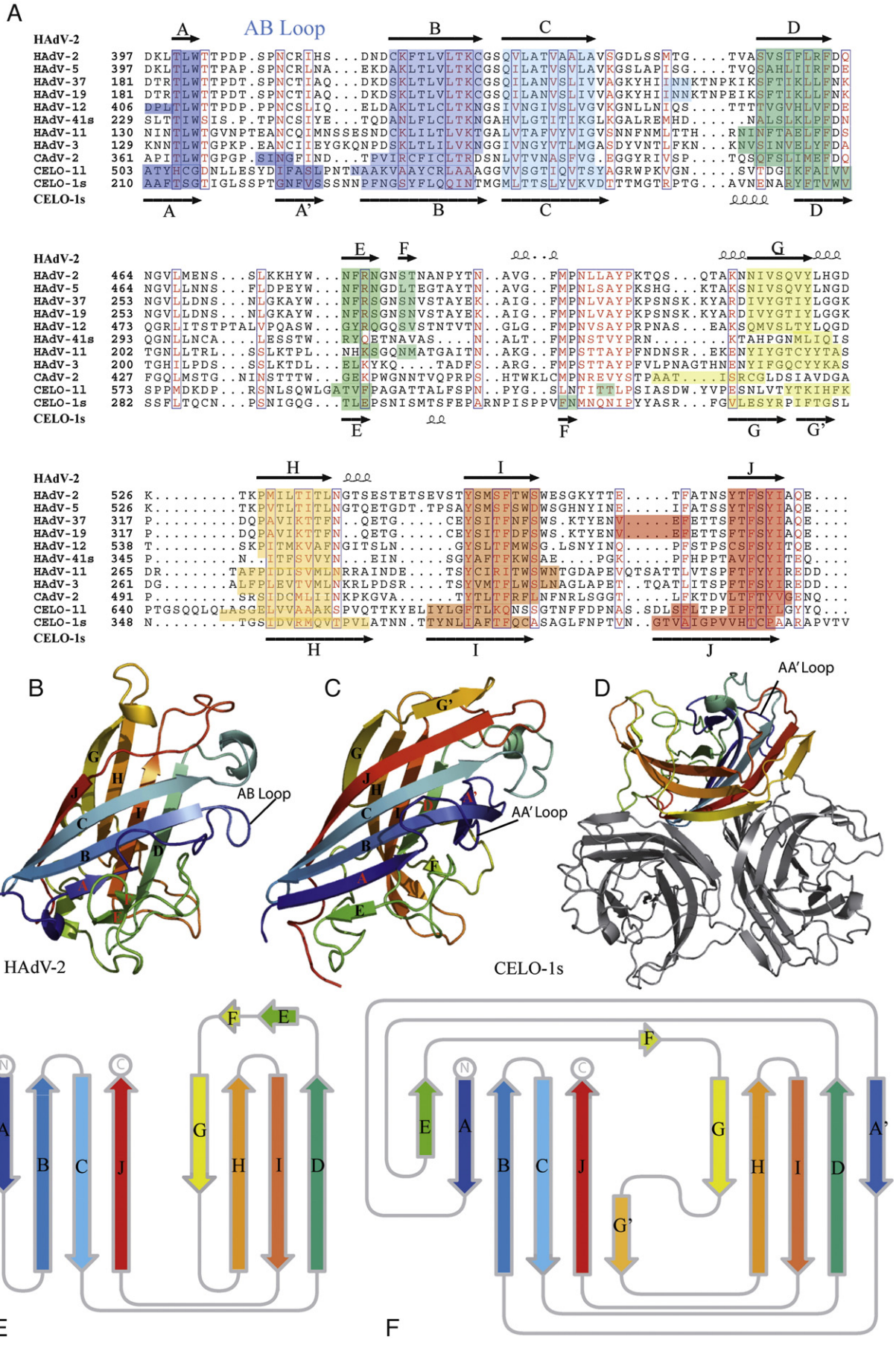
Adenoviruses are non-enveloped, icosahedral viruses, with a linear, 26–45 kbp, double-stranded DNA molecule (Horwitz, 1996; Rekosh et al., 1977; Stewart et al., 1993). Whereas adenoviruses can be found infecting all 5 major vertebrate classes (Davison et al., 2003) only Aviadenovirus that infects birds and Mastadenovirus that infect mammals have been well studied. Avian (fowl) adenovirus 1 or CELO (chicken embryo lethal orphan) virus was isolated in 1957 (Yates and Fry, 1957) and is responsible for mild respiratory pathologies in birds (Dubose and Grumbles, 1959). CELO virus can be isolated from healthy chickens, it does not cause visible disease when it is experimentally introduced into chickens (Cowen et al., 1978) and has not been associated with major economic losses or pathologies in chickens. Most studies on CELO virus have focussed on its genome sequence and organisation whereas virus–cell and virus–host interactions are still poorly understood and structural work on CELO virus proteins has only recently started (Guardado-Calvo et al., 2007; Xu et al., 2007).

The CELO virus genome, 43,804 bp in length, was sequenced (Chiocca et al., 1996) and its transcriptional organization established (Payet et al., 1998). Like the human adenoviruses, CELO virus possesses an icosahedral capsid with a pseudo T number of 25 and a diameter of about 90 nm. Adenovirus capsids are built with 240 copies of the hexon, a homo-trimeric protein that forms the 20 facets of the icosahedron and from 12 pentons that form the 12 vertices (Fabry et al., 2005; Stewart et al., 1991, 1993). The X-ray structure of the CELO virus hexon has been solved at 3.9 Å resolution and has a fold that is similar to human hexons (Xu et al., 2007). For most human and animal adenoviruses, the penton is a non-covalent complex between the pentameric penton base and a single trimeric fibre protein (Fuschiotti et al., 2006; Zubieta et al., 2005). The fibre is a homo-trimer with three distinct regions: the head, the shaft and the tail. The fibre binds to the penton by its N-terminal tail and to the cellular receptor with the head domain. In contrast to other adenoviruses, early observations of CELO virus by electron microscopy revealed pentons with two fibres (Gelderblom and Maichle-Lauppe, 1982; Hess et al., 1995; Laver et al., 1971). The two fibres are of different lengths: the short fibre (fibre 2) is about 8.5 nm long for 410 amino acids and extends straight from the top of the penton base whereas the bottom part of the long fibre (42.5 nm and 710 amino acids in length) lies almost parallel to the top of the penton base (Hess et al., 1995).

The structure of the long-fibre head (amino acids 496–710) was recently solved by X-ray crystallography (Guardado-Calvo et al., 2007). It shows the same over-all fold as the other known adenovirus fibre

* Corresponding authors. R.W.H. Ruigrok is to be contacted at UVHCI, UMR 5233 UJF-EMBL-CNRS, CIBB, BP 181, 38042 Grenoble cedex 9, France. Fax: +33 4 76 20 94 00. G. Schoehn, fax: +33 4 76 20 94 00.
E-mail addresses: ruigrok@embl.fr (R.W.H. Ruigrok), schoehn@embl.fr (G. Schoehn).
[1] Present address: The Division of Structural Biology, Henry Wellcome Building for Genomic Medicine. Roosevelt Drive, Oxford, OX3 7BN, UK.

0042-6822/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2008.05.011


A
AB Loop
HAdV-2
HAdV-5
HAdV-37
HAdV-19
HAdV-12
HAdV-41s
HAdV-11
HAdV-3
CAdV-2
CELO-11
CELO-1s
B
C
D
AB Loop
AA' Loop
HAdV-2
CELO-1s
E
F

heads such as human adenovirus type 5 fibre head (Xia et al., 1994), type 2 (van Raaij et al., 1999), type 12 (Bewley et al., 1999), type 3 (Durmort et al., 2001), types 37 and 19p (Burmeister et al., 2004) and the type 41 short fibre head (Seiradake and Cusack, 2005) and the canine adenovirus type 2 fibre head (Seiradake et al., 2006). Each monomer of the trimeric head is formed by an 8–9 stranded antiparallel β-sandwich connected by loops. The main differences between the various fibre head structures are found in the loop regions that connect the β-strands and in the definition of one or two of the shorter strands. Even though the structures of all fibre heads are very similar, there is very little sequence similarity between the heads of the human adenovirus fibres and even less between fibres of human and animal viruses. The long and short CELO fibre heads show only 15% sequence identity.

Here we have determined the atomic structure of the short-fibre head of CELO (amino acids 201–410) at 2.0 Å resolution. We have used this structure and that of the long-fibre head for a structure based alignment of 11 adenovirus fibre heads that was subsequently used for the construction of an evolutionary tree. Interestingly, both the fibre head sequence and structural alignments suggest that human group F adenovirus 41 (short fibre) is closer to the CELO fibre heads than the canine CAdV-2 fibre head that lies closer to the other human virus fibre heads.

Results and discussion

The head of the short fibre of CELO virus was cloned and expressed in *E. coli*. The protein was crystallised in 0.1 M lithium sulphate, 0.1 M tri-sodium citrate pH 5.6 and 12% w/v polyethylene glycol (PEG) 4000 and its structure solved by isomorphic replacement using a Pt compound. The structure resembles that of all other known structures of adenovirus fibre heads (Fig. 1). Each monomer of the trimer folds into an anti parallel beta sandwich made up by two β-sheets which pack together at an angle of about 45°.

Comparison of the fibre head structures of CELO virus (short fibre) and HAdV2

The nomenclature proposed by Xia et al. (1994) for the fibre head of HAdV5 can also be applied to the short CELO virus fibre head: the beta sandwich is composed of sheets ABCJ and GHID (Fig. 1). The major differences between the CELO and HAdV2/5 heads are: 1; the CELO virus A- and J-strands are longer than the corresponding HAdV2 strands, allowing the A-strand to pack to the E-strand and the J-strand to the G′-strand, see below. 2; the HAdV2 AB loop is interrupted by β-strand A′ in the CELO virus fibre head. 3; strand G is interrupted by a Proline residue in the CELO virus head resulting in two strands, G and G′ where strand G occupies the same position as in the HAdV2 head but where strand G′ contributes to the opposite ABCJ sheet. Point 1 plus point 3 result in a β-sheet that does not have 4 strands but 6: EABCJG′. The extra strand A′ packs onto strand D in the opposite sheet, which, therefore, is composed of 5 strands and not of 4 as in the HAd2 fibre head. The overall effect of these differences is that the CELO fibre head is taller with wider sheets but with shorter loops. The fact that the monomer is taller does not lead to a more compact trimer. The percentage of interface buried upon trimerization is 27 for HAdV2 and only 21 for the CELO short fibre head, due to the fact that the bottom of the CELO fibre head spreads out and sheets A and E hardly contribute to the trimer interface. The main difference between the structures of the heads of the long and short CELO fibres is that in the long fibre head strand A′ does not cross over to the GHIJ sheet but continues in the same direction next to strand B (Guardado-Calvo et al., 2007). This is due to the fact that the B-strand is more strongly curved. In contrast to the long fibre J-strand is broken up in two parts that are separated by a TPP sequence that continues in the same direction as the second part of the J strand and is also in a rather extended conformation, the short fibre J-strand is uninterrupted (Guardado-Calvo et al., 2007).

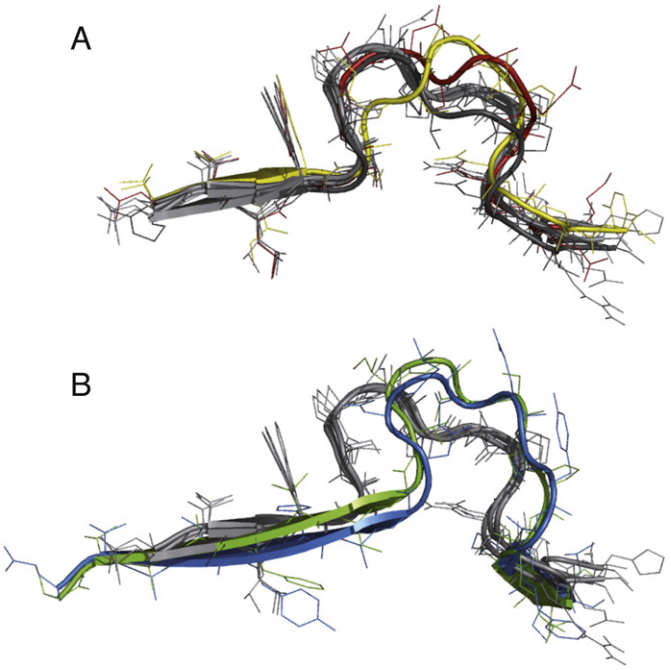


Fig. 2. Comparison of AB loop-structures of CAR-binding and non CAR-binding fibre heads. Ribbon and line representations of the AB loops after structural superposition of different adenovirus fibre head models with Coot (Emsley and Cowtan, 2004). A, AB loops of HAdV-2, 5, 12 (light grey) ((1QHV) (van Raaij et al., 1999) ; (1KNB) (Xia et al., 1994) ; (1KAC) (Bewley et al., 1999)), HadV-37 (dark grey) ((2J12) (Seiradake et al., 2006), HAdV-41s (yellow) ((2BZV) (Seiradake and Cusack, 2005)) and HAdV-11 (red) ((2O39) (Persson et al., 2007)). Note that the conformation of HAdV-37 AB loop corresponds to that found in complex with CAR D1. B, AB loops of HAdV-2, 5, 12 (light grey). AA′ loops of the CELO short (green) and CELO long fibres (blue) ((1IUM) (Guardado-Calvo et al., 2007)).

Receptor binding

The various human adenoviruses bind to different primary receptors. The best known receptor is CAR (coxackie and adenovirus receptor) (Bergelson et al., 1997) that is used by subtypes 2, 5, 12 and 41 long fibre (list non-exhaustive) (Bergelson et al., 1997; Bewley et al., 1999; Roelvink et al., 1998; Tomko et al., 1997). Other human viruses use CD46 (subtypes 3, 7, 11 and 35) (Fleischli et al., 2007; Gaggar et al.,

Fig. 1. Adenovirus fibre head sequences and structures. A, Sequence alignment of all adenovirus fibre heads with known crystal structures: HAdV-2 (van Raaij et al., 1999); HAdV-5 (Xia et al., 1994); HAdV-12 (Bewley et al., 1999) ; HAdV-11 (Persson et al., 2007) ; HAdV-3 (Durmort et al., 2001) ; HAdV-37, HAdV-19p (Burmeister et al., 2004) ; HAdV-41s short (Seiradake and Cusack, 2005); CAdV-2 (Seiradake et al., 2006) ; CELO-1l long (Guardado-Calvo et al., 2007) and CELO-1s short fibre head. Sequences were aligned with ClustalW (Thompson et al., 1994), edited manually with SEAVIEW (Galtier et al., 1996) and visualized with ESPript (Gouet et al., 1999). B, Ribbon diagram of a monomer of the HAdV-2 fibre head, rainbow-coloured from blue (N terminus) to red (C terminus) made with PyMOL (DeLano, 2002). The β-strands are labelled according to the nomenclature introduced by Xia et al. (Xia et al., 1994). **C**, Ribbon diagram of a monomer of the CELO short fibre head, colours as in B. D, Ribbon diagram of CELO short fibre head trimer looking along the 3-fold axis toward the virus. Two monomers are coloured in grey, the third monomer is rainbow-coloured, like in C. E, Strand diagram for the HAdV-2 fibre head monomer (TopDaw, (Bond, 2003)). The colour codes are the same as the one used in part B. F, Strand diagram for CELO short fibre head (TopDaw (Bond, 2003)). The colour codes are the same as the one used in part C.

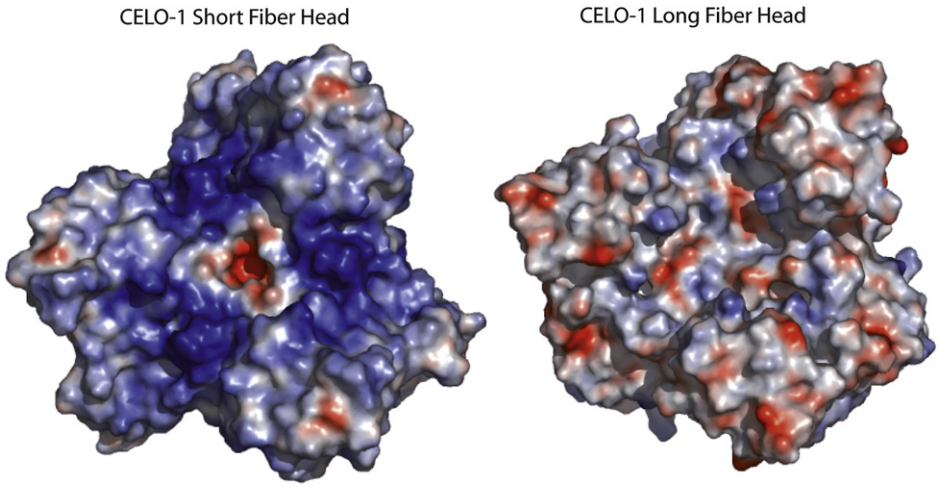


Fig. 3. Comparison of the surface potentials of the CELO short (left) and long (right) fibre heads. Colours are according to the calculated electrostatic surface potential and range from red (potential of −5 kT) to blue (+5 kT). The molecular surface was generated using PyMOL (DeLano, 2002) and electrostatic surface potential was calculated using APBS (Holst et al., 1994).

2003; Persson et al., 2007; Segerman et al., 2003; Sirena et al., 2004) or sialic acid (subtypes 19 and 37) (Burmeister et al., 2004) as receptors. In order to be precise; the recombinant HAd37 fiber knob but also the intact virus bind to CAR *in vitro* but *in vivo* this virus does not use CAR to enter host cells (Wu et al., 2001). Concerning CELO virus, a mutagenesis study showed that the short fibre is essential for infecting chicken cells although the avian receptor responsible for binding of this short fibre is still unknown (Tan et al., 2001). CELO virus could infect CAR expressing mammalian cells but for this the presence of the long-fibre was essential (Tan et al., 2001). However, Guardado-Calvo and coworkers where unable to show *in vitro* binding between the long-fibre head and domain 1 of CAR by surface resonance experiments (Guardado-Calvo et al., 2007). In the HAdV types 2 and 5 the loop most important for CAR binding is the AB loop. Seiradake and Cusack (2005) have suggested that the fibre heads that bind to CAR have a specific conformation of the AB loop. Fig. 2A shows the structures of the AB loops of HAdV2, 5, 12 (in light grey) and 37 (in dark grey) that all bind CAR. The conformation of the HAdV37 loop is that of the CAR-bound structure (Seiradake et al., 2006). Also shown in Fig. 2A are the loops of HAdV41 short fibre (yellow) and that of HAdV11 (red) that do not bind CAR (Fleischli et al., 2007; Roelvink et al., 1998) and have distinctly different conformations.

Fig. 2B again shows the HAdV2, 5 and 12 AB loops (in grey) plus the AA′ loops of the CELO short (green) and long fibres (blue). Due to the fact that the CELO A strand is longer and to the presence of the extra A′ strand, the AA′ loop is squeezed between the two strands and adopts a totally different conformation from that of the CAR-binding loops. This difference can also be seen in the images of the monomers of HAdV2 and CELO in Figs. 1B and C. Therefore, it seems unlikely that the long or short fibres would bind to CAR in the same manner as human adenoviruses types 2, 5 and 12.

Fig. 3 shows the surface charges of the short and long fibres of CELO. The long-fibre head has a slightly acidic to neutral top surface, corresponding to an isoelectric point (p*I*) of 6.1. The top surface of the short-fibre head is more varied in charge with strongly positively charged patches at the monomer interfaces, the pI of this head is 8.5. These differences in surface charges suggest that the two fibres bind to different receptors. Because of the positive charges on the short fibre head we have tested the binding of sialic acid to both fibre heads. Sialic acid is used as receptor by subgroup D adenoviruses (serotypes 8, 15, and 19 and 37) (Arnberg et al., 2000; Burmeister et al., 2004) whose fibre heads all have high p*I* values of around 9. However, we were not able to show binding of either of the CELO fibre heads to sialic acid-containing glycophorin by surface plasma resonance (data not shown).

Evolution

We compared the avian fibre head structure with that of the fibre head of bacteriophage PRD1 (Merckel et al., 2005). It has been suggested that adenovirus and PRD1 fibre heads may be evolutionary related but that a characteristic difference between the adenovirus and bacteriophage fibre heads is that both the adenovirus sheets contain 4 strands whereas the PRD1 sheets contain 5 strands (Bamford et al., 2005; Merckel et al., 2005). Because the CELO heads contain sheets

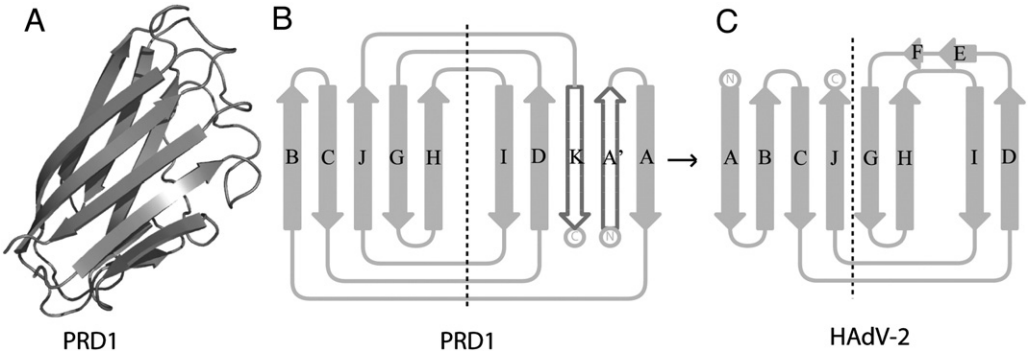


Fig. 4. Structure comparison of the HAdV-2 fibre head and PRD1 vertex protein P5. A, Ribbon diagram of the P5 monomer ((1YQ8) (Merckel et al., 2005)). B, Strand diagram for the PRD1 vertex protein P5. C, Strand diagram for the HAdV-2 fibre head ((1QHV) (van Raaij et al., 1999)) (Figure A was made with PyMOL (DeLano, 2002) and figures B and C were made with TopDaw, (Bond, 2003)).

with 5 and 6 strands, such difference is not characteristic. Fig. 4 shows the structure of a monomer of the PRD1 head plus the strand diagram for the PRD1 and HAdV2 heads (adapted from Fig. 5 in Merckel et al. (2005). These authors have suggested that deletion of the first (A′) and the last strand (K) in the PRD1 head and shift-over of the A strand to the B-strand could explain the difference between the two heads. Comparison of Figs. 4B and C with Fig. 1F clearly shows that the adenovirus heads are much closer in structure to each other than to the bacteriophage head and that the extra strands per sheet have evolved differently for the PRD1 and the CELO heads. Another characteristic difference between the human adenovirus and the bacteriophage head is that the PRD1 sheets cross at an angle between 25 and 30° whereas

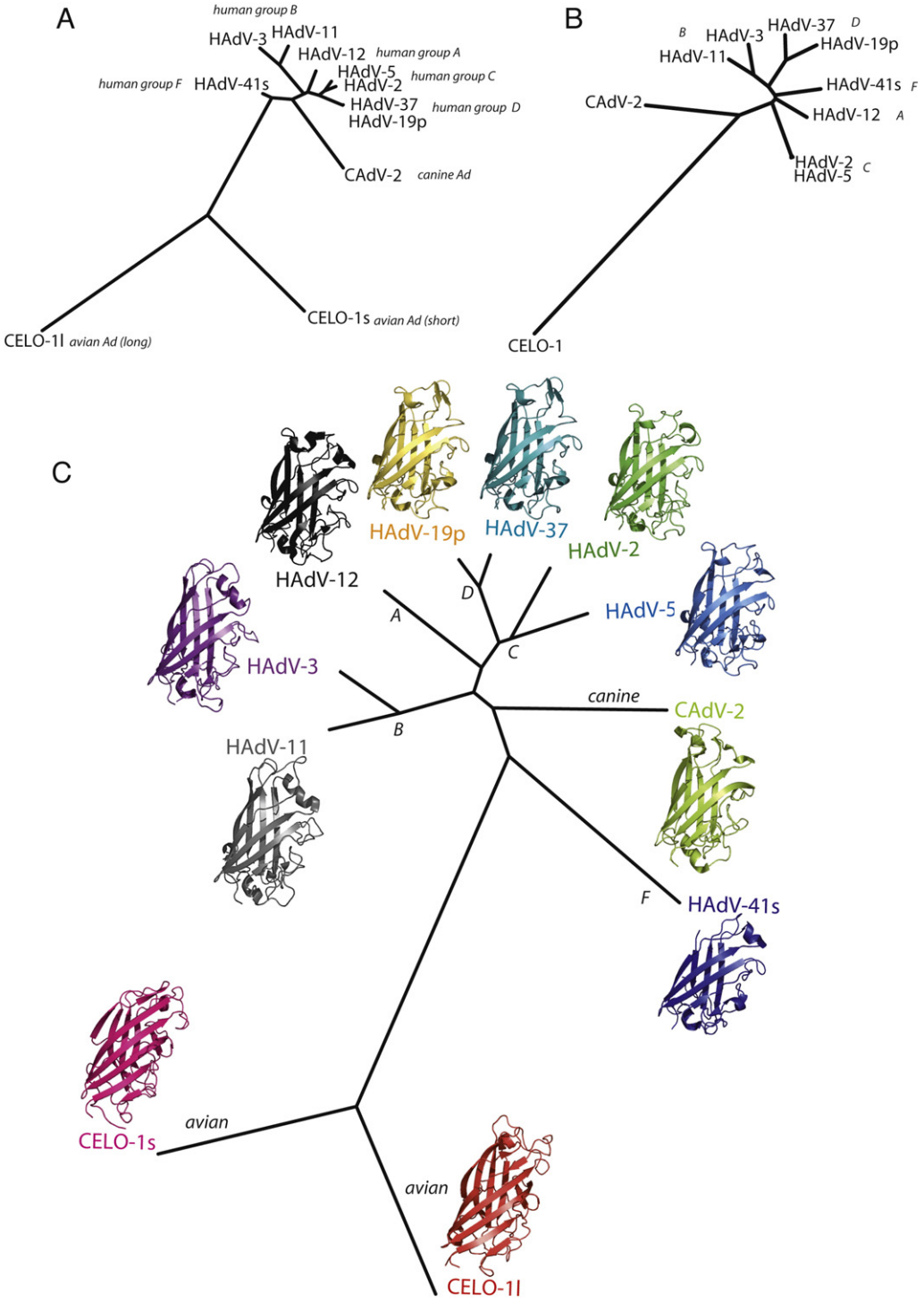


Fig. 5. Phylogenetic trees of adenovirus proteins. A, Distance tree based on Ad fibre head sequences. Protein sequences were aligned as in Fig. 1. The tree was generated with the PHYLIP package (Felsenstein, 1989; Felsenstein, 1997). B, Distance tree based on penton base sequences (PHYLIP). C, Phylogenetic tree of adenovirus fibre head domains based solely on structural comparison using MAPS (Lu, 2000). The PHYLIP package (Felsenstein, 1989, 1997) was used to draw the tree. We compared 11 fibre head structures available from the protein data bank: HAdV-2 (1QHV) (van Raaij et al., 1999) ; HAdV-5 (1KNB) (Xia et al., 1994) ; HAdV-12 (1KAC) (Bewley et al., 1999) ; HAdV-11 (2O39) (Persson et al., 2007) ; HAdV-3 (1H7Z) (Durmort et al., 2001) ; HAdV-37 (1UXE); HAdV-19p (1UXB) (Burmeister et al., 2004) ; HAdV-41s short fibre head (2BZV) (Seiradake and Cusack, 2005); CAdV-2 (2J2J) (Seiradake et al., 2006) ; CELO-1l long fibre head (1IUM) (Guardado-Calvo et al., 2007) and CELO-1s short fibre head.

Table 1
Sequence identity and root mean square deviations (RMSD) between different Ad fibre heads as calculated by Secondary Structure Matching (Krissinel and Henrick, 2004)

Sequence identity (%)											
	HAdV2	HAdV5	HAdV37	HAdV19	HAdV12	HAdV41s	HAdV11	HAdV3	CAdV2	CELO-1l	CELO-1s
HAdV2		66.7	55.8	55.8	40.8	32.5	37.5	38.3	23.3	16.7	8.3
HAdV5	1.43		55.8	55.8	46.7	33.3	33.3	34.2	22.5	15	10
HAdV37	1.558	1.158		99.2	46.7	29.2	37.5	36.7	26.7	14.2	12.5
HAdV19	1.498	1.138	0.321		46.7	29.2	37.5	36.7	26.7	14.2	12.5
HAdV12	1.579	1.551	1.468	1.427		43.3	30.0	36.7	25.8	14.2	10
HAdV41s	1.931	1.777	1.889	1.847	1.733		31.7	30	25.0	14.2	10.8
HAdV11	2.252	1.886	1.728	1.714	1.965	2.292		54.2	21.7	13.3	15.8
HAdV3	2.034	1.772	1.707	1.692	1.886	2.184	1.152		20.8	14.2	10.8
CAdV2	2.248	2.127	2.002	1.929	2.119	2.450	2.297	2.322		15	14.2
CELO-1l	2.907	2.863	2.566	2.503	2.653	2.891	2.834	2.648	2.672		15
CELO-1s	2.835	2.825	2.726	2.638	2.865	3.003	3.132	3.011	3.149	2.607	

RMSD (Å).

the adenovirus sheets cross at about 45°. Here as well the CELO head is characteristic of the adenoviruses crossing at about 50°.

Adenovirus phylogenesis shows co-evolution of the viral subtypes with their host species (Davison et al., 2003). The evolutionary trees are easily visualised using conserved proteins like hexon (Davison et al., 2003) penton base (Madisch et al., 2007) or non-capsid proteins. Comparison of the trees derived from hexons and from pentons suggests recombination events between adenovirus genomes (Madisch et al., 2007). However, the loops on the hexons, on the pentons and virtually the entire fibre are subject to antigenic pressure and show a strong accumulation of mutations. This is the reason why these parts of the capsid are considered to be less suitable for the visualisation of the evolutionary relationships between the viruses. The variability of the fibre sequences could lead to the lining up of β-strand regions in one head with loop regions in other heads.

Table 2
Crystallographic data collection and refinement statistics

Data collection		
	Native	PIP
Detector	ADSC Q315r	ADSC Q315r
Space group	$P4_12_12$	$P4_12_12$
Unit cell (a, b, c, Å)	235.75, 235.75, 61.75	233.81, 233.81, 61.85
(α,β,γ, °)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Wavelength (Å)	0.933	0.933
Resolution	47.6–2.0 (2.10–2.00)	47.6–2.5 (2.65–2.50)
Unique reflections	117179 (15721)	112908 (18068)
Multiplicity	3.63 (3.61)	3.8 (3.8)
Completeness (%)	99.8 (100.0)	99.6 (99.5)
R-merge	12.1 (45.6)	0.118 (0.506)
I/sigma	8.99 (3.19)	8.95 (3.30)
SAD Phasing (SHARP)		
Resolution range	47.4–2.8	
Number heavy atoms	12 Pt	
Phasing power (acentric)	1.14	
Rcullis (acentric)	0.841	
FOM (centric/acentric)	0.156/0.298	
Refinement (REFMAC)		
Resolution	47.6–2.0 (2.05–2.00)	
Reflections used	111098 (7938)	
Reflections Rfree	5865 (402)	
R-factor	0.177 (0.228)	
R-free	0.214 (0.281)	
No. of protein heavy atoms	9245 (2 trimers)	
No. water molecules	1178	
sulphate, glycerol	10, 2	
Average B value	15.8 Å^2	
RMS deviation bond/angles	0.015/1.487	
Ramachandran plot (Molprobity)		
Favored regions	97.8%	
Allowed regions	100%	

Therefore it has been suggested that the alignment of protein structures should tell us more about phylogenetic relationships between viruses than just amino acid sequences (Bamford et al., 2005). We made phylogenetic trees of adenovirus subtypes based both on sequence alignment of penton base and fibre head sequences and a tree based on the structures of the fibre heads. For these alignments we have only used the subtypes for which the head structures were known.

Table 1 shows the amount of sequence identity and the root mean square deviations (RMSD) of the structures for all known head structures. As could be expected, the higher the sequence identity between a pair of heads, the lower the RMSD. The most deviating heads are the two CELO virus heads, the short fibre head being most different from the human virus fibre heads. The canine virus fibre head is much closer to the human than to the avian virus heads.

Fig. 5 shows phylogenetic trees based on alignments of the penton base sequences (Fig. 5B) and on the fibre heads (Fig. 5A) of all viruses for which the head structures are known. The tree based on the hexon sequences is very similar to the penton base tree (data not shown). The main difference between the "head" and the "penton base" trees are that group C viruses are closest to the animal viruses in the penton base tree whereas HAdV41 (group F virus) is closest to the animal viruses in the head tree. In fact, the HadV41 short fibre head sequence locates closer to the CELO head sequence than the canine virus sequence. Apart from the length of the branches, the tree based on the structures of the fibre heads (Fig. 5C) is the same as that based on the head sequences. The two CELO fibre heads have clearly evolved from a common ancestor. If the evolution of mammalian and avian virus sequences have proceeded at similar rates, the duplication of the avian virus fibre sequence must have happened before the human and dog virus lineages separated. In the absence of more mammalian virus head structures it is not possible to say whether all mammalian virus head structure cluster close together or that a transmission occurred from humans to dogs, as is the case for the exclusive human virus that is present in a high percentage of pet dogs (Chiou et al., 2005). It is also possible that a recombination occurred in which a canine virus obtained a human virus fibre or fibre head.

Materials and methods

Protein production

A DNA-fragment encoding residues 201 to 410 of the avian adenovirus short fibre (accession number Q64762) was produced by PCR using primers 5′-CTG ATG GGA TCC CTG TAC CAA GCG CCC ACT AG-3′ (forward primer) and 5′-CTG ATG CTC GAG TCA GAC CGT AAC GGG GGC G-3′ (reverse primer). The PCR product was then cloned between the XhoI and BamH1 restriction sites of the expression vector pProEx HTb (Life Technologies). The resultant plasmid encodes the short fibre fragment fused to an N-terminal purification six-histidine tag.

Protein was expressed over-night at 20 °C in transformed *E. coli* Bl21 star cells (Invitrogen) in Luria Broth medium supplemented with 100 μg/ml ampicilline and 34 μg/ml chloramphenicol. Expression was induced at OD_{600} 0.6 by adding 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation (7 min at 7000 ×g), resuspended in lysis buffer (20 mM Tris–HCl pH 7.9, 500 mM NaCl, 10 mM imidazole, complete EDTA-free protease inhibitor cocktail (Roche) and 0.1 mM PMSF and lysed by sonication. The cell lysate was centrifuged for 45 min at 35,000 ×g and the supernatant containing soluble protein was incubated with cobalt-agarose resin (Sigma) in lysis buffer at 4 °C for 1 h. The resin was washed with the same lysis buffer and then with wash buffer (20 mM Tris–HCl pH 7.9, 500 mM NaCl and 20 mM imidazole) to remove non-specifically bound proteins. The His-tagged short-fibre head protein was then eluted with elution buffer (20 mM Tris–HCl pH 7.9, 200 mM NaCl and 250 mM imidazole). The eluted protein was concentrated and purified by gel filtration on a Superdex 200 column (Amersham Pharmacia Biotech), changing the elution buffer to crystallisation buffer (20 mM Tris–HCl pH 7.9, 200 mM NaCl). Protein homogeneity was checked by SDS-PAGE, negative stain electron microscopy and dynamic light scattering (DLS) (not shown). The protein was concentrated up to 30 mg/ml using a Centricon concentrator (Milipore) and used with the His-tag for crystallization trials.

Crystallogenesis and structure determination

Initial conditions for crystallization of the short-fibre head were identified using a sitting-drop vapour-diffusion screen of commercial sparse-matrix crystallization conditions. The crystallization trials were set up with a Cartesian PixSys 4200 crystallization robot (Genomic Solutions, U.K.) using Greiner Crystal Quick plates (flat bottom, untreated). After optimisation, the best crystals were found to grow at 4 °C in 0.1 M lithium sulphate, 0.1 M tri-sodium citrate at pH 5.6 and 12% w/v polyethylene glycol (PEG) 4000. The crystals were soaked for 5–10 s in a cryoprotectant solution of crystallization buffer plus 25% glycerol. The soaked crystals were mounted on cryo-loops (Hampton), flash frozen at 100 K and placed in a stream of nitrogen at 100 K for all diffraction experiments.

Data were collected on European Synchrotron Radiation Facility (ESRF), beam line ID14-EH4 with an ADSC Q315r CCD detector. The crystals contained two trimers in the asymmetric unit (space group $P4_12_12$), with unit-cell parameters a=235.75 Å, b=235.75 Å, c=61.75 Å and $\alpha=\gamma=\beta=90°$. The best native dataset extended to a resolution of 2.0 Å. All data were integrated and scaled with *XDS* (Kabsch, 1988).

As described by Guardado-Calvo et al. (Guardado-Calvo et al., 2006) for the CELO virus long-fibre head, molecular replacement phasing using human adenovirus fibre heads was not successful, presumably due to very low sequence homology and comparably low structural homology. Therefore, crystals were soaked over night in 6 mM PIP (Di-μ-iodobis(ethylenediamine) diplatinum) and the best derivative crystal datasets extended to 2.5 Å resolution. SHELXD (Sheldrick, 2008) found 12 Pt sites in the asymmetric unit. Inspection showed that these corresponded to two trimers in the asymmetric unit and hence a solvent content of 65%. SHARP (De La Fortelle and Bricogne, 1997) was used to refine the sites and to calculate the initial phases. The map was improved with RESOLVE (Terwilliger, 2002) using initially 2-fold non-crystallographic (NCS) averaging (i.e. the two trimers) and subsequently 6-fold NCS (the six monomers). This allowed the construction of a first polyalanine model for ~80% of the protomer in the asymmetric unit. This model was used for molecular replacement with the 2 Å resolution native dataset with PhaserMR (McCoy et al., 2005). ARP/wARP (Perrakis et al., 1999) was then able to build 1120 residues out of 1422 residues in the asymmetric unit. Model visualization and subsequent building was done with Coot (Emsley and Cowtan, 2004) and refinement with REFMAC (Murshudov et al., 1997). Water molecules were added with ARP/wARP. 10 prominent electron density peaks were assigned to sulphate ions, a component present in the crystallization solution. The final model contains residues 206–410 of each monomer. Density for the five N-terminal amino acids 201–205 was missing from the map. This region is probably flexible or unfolded when the fibre shaft is not complete. The geometry of the final model was checked with MolProbity (Davis et al., 2007) and 98% of all residues were in favoured regions and the remaining 2% were in allowed regions of the Ramachandran plot. The data-collection and refinement statistics are given in Table 2. The coordinates have been deposited in the protein structure database (http://www.rcsb.org) under the accession number 2VTW. The structure factors, including those for the derivative data have also been deposited.

Acknowledgments

We acknowledge the use of the facilities provided through the "Partnership for Structural Biology" such as the crystallization robot; the ESRF for use of beam-line ID14-EH4, Raimond Ravelli for help with data collection and initial data treatment and Hassan Belrhali for help with data collection. GS was supported in part by a Jeunes Chercheurs grant from the Agence Nationale pour la Recherche. MEB is a PhD student financed by the region Rhone Alpes. ES was supported by an "E-STAR" fellowship funded by the EC's FP6 Marie Curie Host fellowship for Early Stage Research Training under contract number MEST-CT-2004-504640.We thank Patrick Langlois, AFSSA, Ploufragan, France for providing the DNA clone with the CELO fibre head sequence.

References

Arnberg, N., Edlund, K., Kidd, A.H., Wadell, G., 2000. Adenovirus type 37 uses sialic acid as a cellular receptor. J Virol 74 (1), 42–48.
Bamford, D.H., Grimes, J.M., Stuart, D.I., 2005. What does structure tell us about virus evolution? Curr. Opin. Struct. Biol. 15 (6), 655–663.
Bergelson, J.M., Cunningham, J.A., Droguett, G., Kurt-Jones, E.A., Krithivas, A., Hong, J.S., Horwitz, M.S., Crowell, R.L., Finberg, R.W., 1997. Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. Science 275 (5304), 1320–1323.
Bewley, M.C., Springer, K., Zhang, Y.B., Freimuth, P., Flanagan, J.M., 1999. Structural analysis of the mechanism of adenovirus binding to its human cellular receptor, CAR. Science 286 (5444), 1579–1583.
Bond, C.S., 2003. TopDraw: a sketchpad for protein structure topology cartoons. Bioinformatics 19 (2), 311–312.
Burmeister, W.P., Guilligay, D., Cusack, S., Wadell, G., Arnberg, N., 2004. Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. J Virol 78 (14), 7727–7736.
Chiocca, S., Kurzbauer, R., Schaffner, G., Baker, A., Mautner, V., Cotten, M., 1996. The complete DNA sequence and genomic organization of the avian adenovirus CELO. J. Virol. 70 (5), 2939–2949.
Chiou, S.H., Chow, K.C., Yang, C.H., Chiang, S.F., Lin, C.H., 2005. Discovery of Epstein–Barr virus (EBV)-encoded RNA signal and EBV nuclear antigen leader protein DNA sequence in pet dogs. J. Gen. Virol. 86 (Pt 4), 899–905.
Cowen, B., Calnek, B.W., Menendez, N.A., Ball, R.F., 1978. Avian adenoviruses: effect on egg production, shell quality, and feed consumption. Avian Dis. 22 (3), 459–470.
Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall III, W.B., Snoeyink, J., Richardson, J.S., Richardson, D.C., 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 35, W375–W383 (Web Server issue).
Davison, A.J., Benko, M., Harrach, B., 2003. Genetic content and evolution of adenoviruses. J. Gen. Virol. 84 (Pt 11), 2895–2908.
De La Fortelle, E., Bricogne, G., 1997. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol. 276, 472–494.
DeLano, W.L., 2002. The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA.
Dubose, R.J., Grumbles, L.C., 1959. The relationship between quail bronchitis virus and chicken embryo lethal orphan virus. Avian Dis. 3, 321–344.
Durmort, C., Stehlin, C., Schoehn, G., Mitraki, A., Drouet, E., Cusack, S., Burmeister, W.P., 2001. Structure of the fiber head of Ad3, a non-CAR-binding serotype of adenovirus. Virology 285 (2), 302–312.
Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D. Biol. Crystallogr. 60 (Pt 12 Pt 1), 2126–2132.
Fabry, C.M., Rosa-Calatrava, M., Conway, J.F., Zubieta, C., Cusack, S., Ruigrok, R.W., Schoehn, G., 2005. A quasi-atomic model of human adenovirus type 5 capsid. EMBO J. 24 (9), 1645–1654.
Felsenstein, J., 1989. Mathematics vs. evolution: mathematical evolutionary theory. Science 246 (4932), 941–942.
Felsenstein, J., 1997. An alternating least squares approach to inferring phylogenies from pairwise distances. Syst Biol. 46 (1), 101–111.

Fleischli, C., Sirena, D., Lesage, G., Havenga, M.J., Cattaneo, R., Greber, U.F., Hemmi, S., 2007. Species B adenovirus serotypes 3, 7, 11 and 35 share similar binding sites on the membrane cofactor protein CD46 receptor. J. Gen. Virol. 88 (Pt 11), 2925–2934.
Fuschiotti, P., Schoehn, G., Fender, P., Fabry, C.M., Hewat, E.A., Chroboczek, J., Ruigrok, R.W., Conway, J.F., 2006. Structure of the dodecahedral penton particle from human adenovirus type 3. J. Mol. Biol. 356 (2), 510–520.
Gaggar, A., Shayakhmetov, D.M., Lieber, A., 2003. CD46 is a cellular receptor for group B adenoviruses. Nat. Med. 9 (11), 1142–1408.
Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12 (6), 543–548.
Gelderblom, H., Maichle-Lauppe, I., 1982. The fibers of fowl adenoviruses. Arch. Virol. 72 (4), 289–298.
Gouet, P., Courcelle, E., Stuart, D.I., Metoz, F., 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15 (4), 305–308.
Guardado-Calvo, P., Llamas-Saiz, A.L., Langlois, P., van Raaij, M.J., 2006. Crystallization of the C-terminal head domain of the avian adenovirus CELO long fibre. Acta Crystallogr., Sect. F Struct. Biol. Cryst. Commun. 62 (Pt 5), 449–452.
Guardado-Calvo, P., Llamas-Saiz, A.L., Fox, G.C., Langlois, P., van Raaij, M.J., 2007. Structure of the C-terminal head domain of the fowl adenovirus type 1 long fiber. J. Gen. Virol. 88 (Pt 9), 2407–2416.
Hess, M., Cuzange, A., Ruigrok, R.W., Chroboczek, J., Jacrot, B., 1995. The avian adenovirus penton: two fibres and one base. J. Mol. Biol. 252 (4), 379–385.
Holst, M., Kozack, R.E., Saied, F., Subramaniam, S., 1994. Protein electrostatics: rapid multigrid-based Newton algorithm for solution of the full nonlinear Poisson–Boltzmann equation. J. Biomol. Struct Dyn. 11 (6), 1437–1445.
Horwitz, M., 1996. Adenoviruses. In: Fields, B., Knipe, D. (Eds.), Fields Virology, vol. 2. Raven Press, Philadelphia, PA, pp. 2149–2171. 2 vols.
Kabsch, W., 1988. Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. J. Appl. Crystallogr. 21, 916–924.
Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr., D Biol. Crystallogr. 60 (Pt 12 Pt 1), 2256–2268.
Laver, W.G., Younghusband, H.B., Wrigley, N.G., 1971. Purification and properties of chick embryo lethal orphan virus (an avian adenovirus). Virology 45 (3), 598–614.
Lu, G., 2000. TOP: a new method for protein structure comparisons and similarity searches. J. Appl. Crystallogr. 33, 176–183.
Madisch, I., Hofmayer, S., Moritz, C., Grintzalis, A., Hainmueller, J., Pring-Akerblom, P., Heim, A., 2007. Phylogenetic analysis and structural predictions of human adenovirus penton proteins as a basis for tissue-specific adenovirus vector design. J. Virol. 81 (15), 8270–8281.
McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., Read, R.J., 2005. Likelihood-enhanced fast translation functions. Acta Crystallogr., D Biol. Crystallogr. 61 (Pt 4), 458–464.
Merckel, M.C., Huiskonen, J.T., Bamford, D.H., Goldman, A., Tuma, R., 2005. The structure of the bacteriophage PRD1 spike sheds light on the evolution of viral capsid architecture. Mol. Cell 18 (2), 161–170.
Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr., D Biol. Crystallogr. 53 (Pt 3), 240–255.
Payet, V., Arnauld, C., Picault, J.P., Jestin, A., Langlois, P., 1998. Transcriptional organization of the avian adenovirus CELO. J. Virol. 72 (11), 9278–9285.
Perrakis, A., Morris, R., Lamzin, V.S., 1999. Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. 6 (5), 458–463.
Persson, B.D., Reiter, D.M., Marttila, M., Mei, Y.F., Casasnovas, J.M., Arnberg, N., Stehle, T., 2007. Adenovirus type 11 binding alters the conformation of its receptor CD46. Nat. Struct. Mol. Biol. 14 (2), 164–166.
Rekosh, D.M., Russell, W.C., Bellet, A.J., Robinson, A.J., 1977. Identification of a protein linked to the ends of adenovirus DNA. Cell 11 (2), 283–295.
Roelvink, P.W., Lizonova, A., Lee, J.G., Li, Y., Bergelson, J.M., Finberg, R.W., Brough, D.E., Kovesdi, I., Wickham, T.J., 1998. The coxsackievirus–adenovirus receptor protein can function as a cellular attachment protein for adenovirus serotypes from subgroups A, C, D, E, and F. J. Virol. 72 (10), 7909–7915.
Segerman, A., Atkinson, J.P., Marttila, M., Dennerquist, V., Wadell, G., Arnberg, N., 2003. Adenovirus type 11 uses CD46 as a cellular receptor. J. Virol. 77 (17), 9183–9191.
Seiradake, E., Cusack, S., 2005. Crystal structure of enteric adenovirus serotype 41 short fiber head. J. Virol. 79 (22), 14088–14094.
Seiradake, E., Lortat-Jacob, H., Billet, O., Kremer, E.J., Cusack, S., 2006. Structural and mutational analysis of human Ad37 and canine adenovirus 2 fiber heads in complex with the D1 domain of coxsackie and adenovirus receptor. J. Biol. Chem. 281 (44), 33704–33716.
Sheldrick, G.M., 2008. A short history of SHELX. Acta Crystallogr., A 64 (Pt 1), 112–122.
Sirena, D., Lilienfeld, B., Eisenhut, M., Kalin, S., Boucke, K., Beerli, R.R., Vogt, L., Ruedl, C., Bachmann, M.F., Greber, U.F., Hemmi, S., 2004. The human membrane cofactor CD46 is a receptor for species B adenovirus serotype 3. J. Virol. 78 (9), 4454–4462.
Stewart, P.L., Burnett, R.M., Cyrklaff, M., Fuller, S.D., 1991. Image reconstruction reveals the complex molecular organization of adenovirus. Cell 67 (1), 145–154.
Stewart, P.L., Fuller, S.D., Burnett, R.M., 1993. Difference imaging of adenovirus: bridging the resolution gap between X-ray crystallography and electron microscopy. EMBO J. 12 (7), 2589–2599.
Tan, P.K., Michou, A.I., Bergelson, J.M., Cotten, M., 2001. Defining CAR as a cellular receptor for the avian adenovirus CELO using a genetic analysis of the two viral fibre proteins. J. Gen. Virol. 82 (Pt 6), 1465–1472.
Terwilliger, T.C., 2002. Statistical density modification with non-crystallographic symmetry. Acta Crystallogr., D Biol. Crystallogr. 58 (Pt 12), 2082–2086.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22 (22), 4673–4680.
Tomko, R.P., Xu, R., Philipson, L., 1997. HCAR and MCAR: the human and mouse cellular receptors for subgroup C adenoviruses and group B coxsackieviruses. Proc. Natl. Acad. Sci.U. S. A. 94 (7), 3352–3356.
van Raaij, M.J., Louis, N., Chroboczek, J., Cusack, S., 1999. Structure of the human adenovirus serotype 2 fiber head domain at 1.5 Å resolution. Virology 262 (2), 333–343.
Wu, E., Fernandez, J., Fleck, S.K., Von Seggern, D.J., Huang, S., Nemerow, G.R., 2001. Virology 279, 78–89.
Xia, D., Henry, L.J., Gerard, R.D., Deisenhofer, J., 1994. Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution. Structure 2 (12), 1259–1270.
Xu, L., Benson, S.D., Burnett, R.M., 2007. Nanoporous crystals of chicken embryo lethal orphan (CELO) adenovirus major coat protein, hexon. J. Struct. Biol. 157 (2), 424–431.
Yates, V.J., Fry, D.E., 1957. Observations on a chicken embryo lethal orphan (CELO) virus. Am. J. Vet. Res. 18 (68), 657–660.
Zubieta, C., Schoehn, G., Chroboczek, J., Cusack, S., 2005. The structure of the human adenovirus 2 penton. Mol. Cell 17 (1), 121–135.