

Crystal Structure of Species D Adenovirus Fiber Knobs and Their Sialic Acid Binding Sites

Wim P. Burmeister,^{1,2*} Delphine Guilligay,² Stephen Cusack,² Göran Wadell,³ and Niklas Arnberg³

Laboratoire de Virologie Moléculaire et Structurale, EA 2939, Université Joseph Fourier, Faculté de Médecine de Grenoble, F-38706 La Tronche,¹ and EMBL, Grenoble Outstation, F-38042 Grenoble cedex 9,² France, and Department of Virology, Umeå University, SE-90185 Umeå, Sweden³

Received 17 January 2004/Accepted 26 February 2004

Adenovirus serotype 37 (Ad37) belongs to species D and can cause epidemic keratoconjunctivitis, whereas the closely related Ad19p does not. Primary cell attachment by adenoviruses is mediated through receptor binding of the knob domain of the fiber protein. The knobs of Ad37 and Ad19p differ at only two positions, Lys240Glu and Asn340Asp. We report the high-resolution crystal structures of the Ad37 and Ad19p knobs, both native and in complex with sialic acid, which has been proposed as a receptor for Ad37. Overall, the Ad37 and Ad19p knobs are very similar to previously reported knob structures, especially to that of Ad5, which binds the coxsackievirus-adenovirus receptor (CAR). Ad37 and Ad19p knobs are structurally identical with the exception of the changed side chains and are structurally most similar to CAR-binding knobs (e.g., that of Ad5) rather than non-CAR-binding knobs (e.g., that of Ad3). The two mutations in Ad19p result in a partial loss of the exceptionally high positive surface charge of the Ad37 knob but do not affect sialic acid binding. This site is located on the top of the trimer and binds both $\alpha(2,3)$ and $\alpha(2,6)$ -linked sialyl-lactose, although only the sialic acid residue makes direct contact. Amino acid alignment suggests that the sialic acid binding site is conserved in several species D serotypes. Our results show that the altered viral tropism and cell binding of Ad19p relative to those of Ad37 are not explained by a different binding ability toward sialyl-lactose.

Adenoviruses form nonenveloped icosahedral particles that contain a double-stranded DNA genome. They are built mainly from 240 copies of hexon proteins and 12 pentons, consisting of trimeric fibers inserted into the pentameric penton bases. Despite a few exceptions, the most commonly used mechanism whereby adenoviruses attach to host cells involves an interaction between a host cell receptor and the knob domain of the viral fiber protein, whereas the penton base is essential for internalization (reviewed in references 8, 44, and 55). The fiber protein is trimeric and is divided into tail, shaft, and knob. The tail anchors the fiber to penton base capsomers, whereas the shaft is built from 6 to 22 copies of a repeating motif. It differs in length and rigidity among species and allows the knob to protrude from the virion and interact with cellular receptors (15).

The 51 characterized human adenoviruses are divided into six different species (A through F [7]). Adenoviruses belonging to species A (adenovirus serotype 12 [Ad12] and Ad31), C (Ad2 and Ad5), D (Ad9 and Ad15), E (Ad4), and F (Ad41) have been shown to interact *in vitro* with a cellular receptor designated CAR (coxsackie-adenovirus receptor) (9, 43, 51). CAR, a member of the immunoglobulin superfamily (9), is able to form intercellular homodimers (52) and appears to be an important component in epithelial tight junctions (17).

Species D contains 32 different serotypes, but most of these serotypes are rarely isolated from humans. However, all isolates of the three serotypes Ad8, Ad19, and Ad37, with the single exception of the prototype isolate of Ad19 (Ad19p), cause a severe ocular disease, epidemic keratoconjunctivitis (EKC) (23, 25).

Different possible receptors for Ad37 have been identified, and their respective roles remain controversial.

On the one hand, Wu et al. (60) showed that Ad37 attached to Chang C cells through a 50-kDa protein in a calcium-dependent and sialic acid-independent manner. Recently, the ligand has been identified as CD46 (59). Human CD46, also known as membrane cofactor protein (MCP), is a widely expressed glycoprotein that is present on most cells. It binds to complement proteins C3b and C4b at the cell surface and is used by some strains of measles virus (19, 38) and human herpesvirus 6 (45) as a cellular receptor. CD46 has been identified recently as a ligand for the species B virus Ad11p (47), Ad35 (24), and other adenoviruses from species B1 (Ad16, Ad21, and Ad50) and B2 (Ad14) (24). The question of the receptor used by Ad3 is not yet settled. Whereas Gaggar et al. (24) found that Ad3 cannot use CD46, Sirena and coworkers found that it is a receptor (48).

On the other hand, it has been demonstrated that all EKC-causing adenoviruses use sialic acid-containing oligosaccharides as cellular receptors instead of CAR (3, 4, 5). This usage matches the unique pathogenicity of the EKC-causing adenoviruses (Ad8, Ad19a, and Ad37). The interaction between these viruses and sialic acid is mediated by a charge-dependent mechanism between the unusually positively charged fiber knobs of these viruses (predicted pI, around 9 [see Fig. 1] [6, 37]) and sialic acid (2). Also, Segerman et al. (47) showed that CHO cells bind Ad37 virus whereas overexpression of CD46 in these cells does not increase binding.

The fiber knob of Ad19p differs in two positions from Ad37: Lys240 and Asn340 for Ad37 are replaced by glutamate and aspartic acid in Ad19p, changing the predicted pI from 9.22 to 8.84. Huang et al. (28) showed that Ad37 binds with high affinity ($K_d = 3.5$ nM) to Chang C cells and that this interaction

* Corresponding author. Mailing address: EMBL, Grenoble Outstation, B.P. 181, F-38042 Grenoble cedex 9, France. Phone: 33 4 76 20 72 82. Fax: 33 4 76 20 71 99. E-mail: wpb@embl-grenoble.fr.

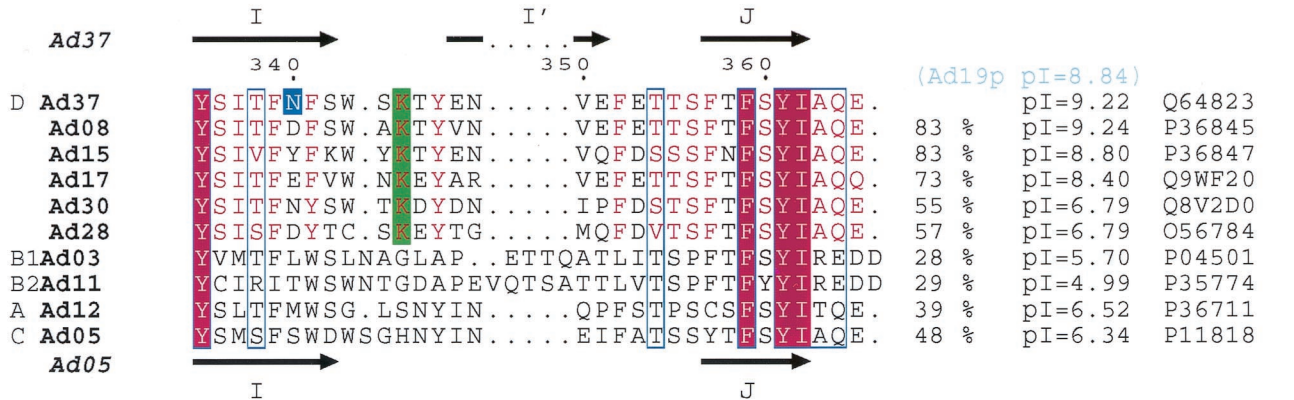
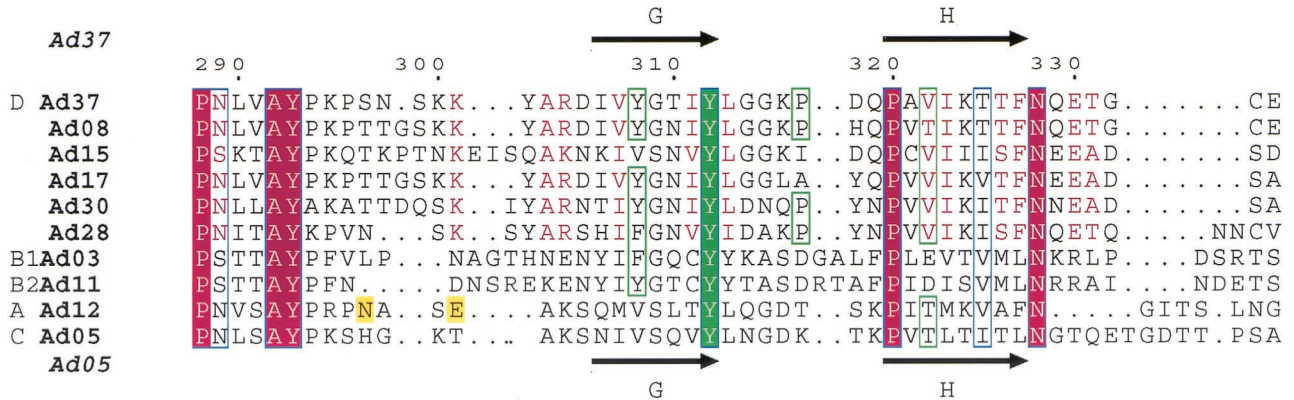
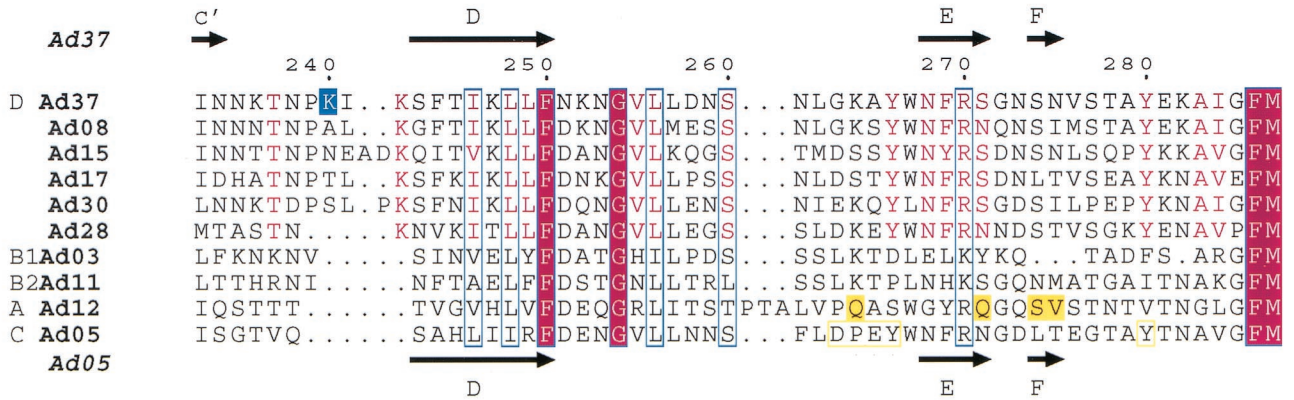
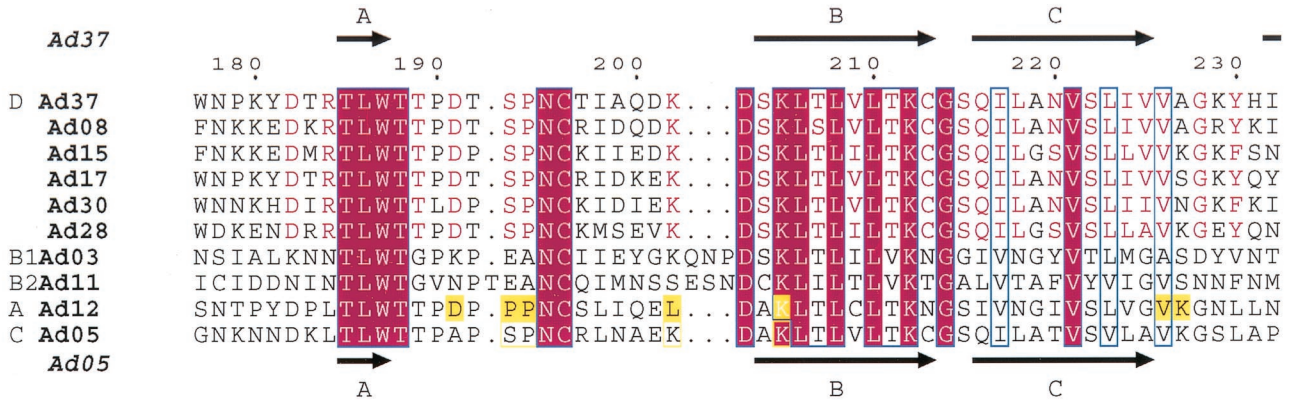


TABLE 1. Statistics on refinement and data collection of the principal data sets and structures

Parameter	Value for the following structure (PDB ^a entry):		
	Ad37 native (1uxc)	Ad37-sialyl-lactose (10 mM) (1uxa)	Ad19p-sialyl-lactose (10 mM) (1uxb)
Data collection			
Resolution (Å)	2.0 (2.11–2.00)	1.5 (1.59–1.50)	1.75 (1.84–1.75)
Completeness	0.982 (0.969)	0.935 (0.666)	0.999 (0.999)
Multiplicity	3.6 (3.3)	3.4 (2.9)	4.3 (3.6)
R_{merge}	0.107 (0.25)	0.081 (0.26)	0.106 (0.34)
Refinement			
No. of reflections	42,597	90,108	61,498
No. of water molecules	385	529	587
R_{cryst}	0.204 (0.286)	0.180 (0.266)	0.175 (0.268)
R_{free}	0.241 (0.339)	0.205 (0.315)	0.210 (0.309)
SIGMAA coordinate error (test set, Å)	0.29	0.17	0.17
rms bonds	0.023	0.023	0.016
rms angles (°)	2.0	2.1	1.9

^a PDB, Protein Data Bank.

determines the tropism of Ad37, as Ad19p does not bind. This could be further pinned down further to the Lys240Glu mutation. The change of tropism has been attributed to the change in the electrostatic surface potential of the fiber knob (2).

The fiber knobs of Ad2, Ad3, Ad5, and Ad12 have been crystallized previously (10, 22, 54, 61). The fiber knob of Ad12 has also been crystallized in complex with CAR (10). The CAR binding site is located at the interface between two adjacent knob monomers. It has been shown that the model of the CAR binding site obtained from the Ad12–CAR complex structure is valid as well for species C (27) and D (31) viruses. Kirby et al. (31) have shown that the Ad9 serotype of species D binds CAR with a micromolar affinity, which, however, is several orders of magnitude lower than the affinities of Ad5 and Ad12.

To understand better the structural mechanisms involved in adenovirus interactions with cellular receptors different from CAR and the reasons for the differences in tropism between Ad37 and Ad19p, we determined the 3-dimensional structures of the Ad37 and Ad19p knobs, both alone and in complex with sialyl-lactose.

MATERIALS AND METHODS

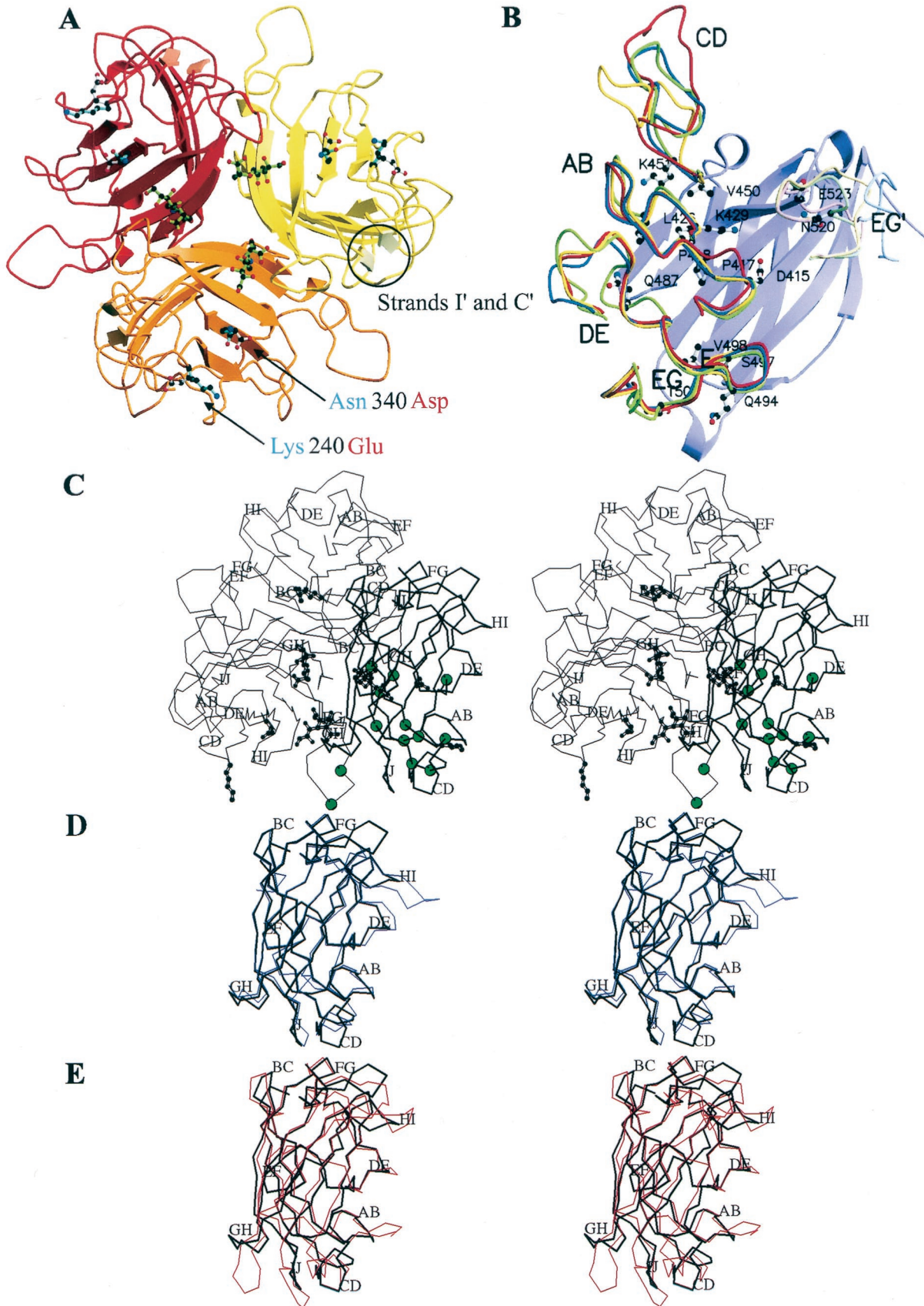
Protein expression and purification. The genes corresponding to residues 177 to 365 of the full-length fibers were amplified from viral DNA by using primers 5'-CGCGGATCTGGAACCCAAAGTAT-3' and 5'-CCGCTCGAGTCATTC TTGGCAAT-3' and were then cloned into the pPROEX HTb plasmid (Life Technologies) by using a BamHI and an XhoI restriction site. The construct has been verified by DNA sequencing. The plasmid contained an N-terminal His tag and a tobacco etch virus (TEV) protease cleavage site and was expressed in BL21 Star (DE3) cells (Life Technologies) at 37°C. The pellet from a 3-liter culture was treated with 1 mg of lysozyme/ml on ice and sonicated 10 times for 10 s each time. After centrifugation at 40,000 × g for 30 min, 10 mM imidazole and 15 ml nitrilotriacetic acid resin (Pharmacia) were added to the soluble fraction at 4°C

with gentle stirring. The resin was filled in a small polyethylene column and washed with 20 mM imidazole in a buffer containing 30 mM Tris-HCl (pH 7.5) and 150 mM NaCl. The protein was eluted with 500 mM imidazole in the same buffer. The His tag was cleaved off by an overnight incubation with 1/100 His-tagged TEV protease (50) at 10°C. Imidazole was eliminated by dialysis against the buffer. Uncleaved fiber knob and TEV protease were bound to 7 ml of nitrilotriacetic acid resin in a buffer containing 20 mM imidazole at 4°C. The supernatant was concentrated, purified over a Superdex 200 column (Pharmacia), and concentrated to 11 mg/ml.

Crystallization and structure determination. Crystals of Ad37 and Ad19p fiber knobs were grown by the hanging drop method using a reservoir solution of 24% polyethylene glycol 8000, 50 mM zinc acetate, and 100 mM HEPES (pH 7.5). For sialic acid binding, crystals were soaked in the reservoir solution containing additional 10 mM sialyl-lactose (from human colostrum; Boehringer Mannheim). Crystals could be frozen prior to data collection in the 100 K nitrogen gas stream without cryoprotection. Data were collected on beamlines ID29, ID14-2, and ID14-3 at the European Synchrotron Radiation Facility (Grenoble, France) by using ADSC Quantum4 area detectors. Images were processed by using MOSFLM (36) and the CCP4 program suite (18). The structure of Ad37 was solved first by molecular replacement using AMORE (39) and the Ad2 knob trimer (54) as a search model. Eighty-nine percent of the residues which were finally visible could be built with ARP/wARP (41). Chain tracing was completed by using O (30), and the structure was refined by using the crystallography and nuclear magnetic resonance system (CNS) (11). Statistics on data collection and refinement are given in Table 1. Crystals of Ad19p differed only very slightly in the cell parameters and were otherwise isomorphous except for the two amino acid mutations differentiating Ad37 and Ad19p. Coordinates have been deposited in the Protein Data Bank.

Calculation of the ligand occupancy. Crystals of Ad19p and Ad37 fiber knobs were soaked for a minimum of 2 h in different concentrations of sialyl-lactose in the reservoir solution. Data sets covering 180° were collected at maximal resolutions (d_{max}) ranging from 1.75 to 2.2 Å and processed as described above. The unit cells of all crystals were treated as identical and isomorphous to the native Ad37–10 mM sialic acid data set. The Ad37 and Ad19p structures determined at a sialyl-lactose concentration of 10 mM were used as reference structures (Table 1). They were subjected to a refinement of the overall temperature factor and a positional refinement (leaving the sialyl-lactose group fixed) against the different data sets. ($F_o - F_c$)-omit maps where the sialyl-lactose atoms and the atoms of tyrosine residues 308 and

FIG. 1. Alignment of adenovirus knob sequences. Group 1 contains representative sequences of viruses of species D, as well as Ad3 and Ad11p (species B), Ad12 (species A), and Ad5 (species C). Secondary-structure elements of Ad37 are shown above the alignment (arrows, beta strands). The secondary structures of Ad5, the closest relative of Ad37 with a previously known structure, is shown below the alignment. Conserved residues are in blue boxes. The two key residues forming the sialic acid binding site and their conserved counterparts are shown on a green background; other residues interacting with sialic acid are boxed in green. Residues contacting CAR in the AD12–CAR complex crystal structure are shown on a yellow background, whereas residues of Ad5 identified as interacting with CAR by mutagenesis (32, 33) are boxed in yellow. A blue background highlights the two residues that are different in the Ad37 versus the Ad19p fiber knob. Residues conserved within species D are printed in red. ESPript software (26) was used to prepare this figure. Sequence identities and calculated pI values are based on the sequence (SPTREMBL entry given) of the knob domain as shown in this figure.



312 have been left out. The electron density belonging to these groups has been estimated by a summation of the values of the grid points calculated on a grid of $d_{\text{max}}/4$ within 1.5 \AA of the atoms of the residues concerned (by using MAMA and MAPMAN software [34, 35]). Based on the number of electrons belonging to sialic acid ($182 e^-$) and tyrosine ($86 e^-$), occupancies of the sialic acid groups can be calculated by using the tyrosine groups as internal standards with an occupancy of 1. Because the tyrosines are part of the sialic acid binding site, temperature factors are likely to be in the same range as for the ligand. Only the two $\alpha(2,3)$ -linked ligands have been considered, since the affinity of the third site is affected by a crystal contact which would not be present in solution. For the Ad37 crystal soaked with 10 mM sialyl-lactose, reinjection of the refined occupancies into a temperature factor refinement yielded an average temperature factor of 20 \AA^2 for sialic acid, in agreement with the temperature factors of the contacting amino acid residues (range, 12 to 22 \AA^2). Assuming a simple equilibrium, the occupancy (q) as a function of the substrate concentration (C) can be calculated as $q = \frac{C}{C + K_d}(1 - q_0) + q_0$, where q_0 is the electron density due to bound water present in the absence of bound ligand. A q_0 of 0.085 has been used for Ad37 (0.15 for Ad19p). Estimates for K_d were obtained by using a least-squares fit to the experimental data points.

RESULTS

The fiber knob domain of Ad37 containing residues 177 to 365 of the native fiber could be crystallized in spacegroup $P2_1$ with the following cell parameters: $a = 60.7$, $b = 68.7$, $c = 74.7 \text{ \AA}$, $\beta = 94.9^\circ$. The asymmetric unit contains one trimer. Crystals that had been soaked with 10 mM sialyl-lactose diffracted to 1.5 \AA resolution and were used to solve the structure by molecular replacement based on the model of the Ad2 fiber knob trimer (54).

In the crystal, the knob trimers interact mainly through zinc ions, identified by a strong electron density next to His231 and different carboxylic acid residues of neighboring molecules. Because His231 does not show any conservation across related sequences (Fig. 1), we consider the zinc binding to be an artifact due to the presence of zinc in the precipitant.

The structure shows the residues from Thr183 to Glu365, the natural C terminus of the protein (Fig. 2A). Overall, the fiber knob closely resembles previously determined knob structures. The central β -sandwich structure as well as the subunit orientation is well conserved. A least-squares superposition of the $C\alpha$ atoms of the different structures (Ad2, 155 atoms, 0.99 \AA rms; Ad3, 150 atoms, 1.36 \AA rms; Ad5, 168 atoms, 0.93 \AA rms; Ad12, 159 atoms, 1.29 \AA rms) reveals that the CAR binding serotype Ad5 is the most closely related structure (Fig. 2B and D). Relative to other fiber knob structures, an additional small β -sheet formed by strands C' and I' (Fig. 1 and 2A) is located on the top of the structure.

By use of the analogous construct, fiber heads from Ad19p crystallized in the same way in crystals with very similar cell parameters. The two structures are identical except for the two amino acid changes in Ad19p: Lys240Glu and Asn340Asp. In

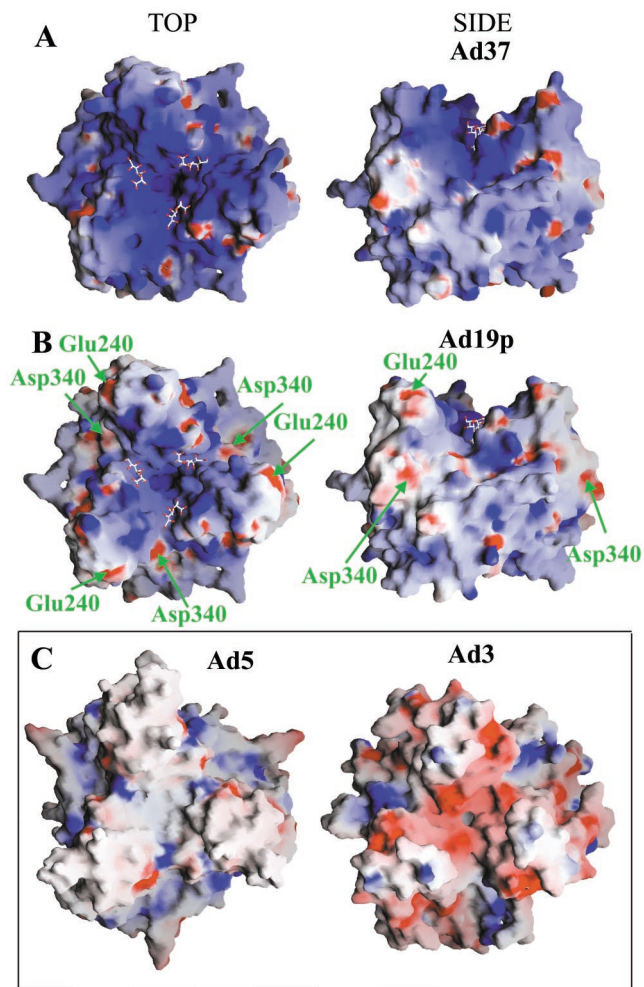


FIG. 3. Electrostatic surface potential of fiber knobs. (A) Top and side views of the Ad37 knob with bound sialic acid residues. (B) Ad19p knob. Locations of its specific mutations relative to Ad37 are shown. (C) Top views of Ad5 and Ad3 knobs in the same orientation. The figure was prepared with GRASP (40). Coloring of the surface potential ranges from $+10 \text{ kT}$ (blue) to -10 kT (red).

Ad37, Lys240 is located on the very top of the fiber knob in the CD loop, where the residue is oriented outward (Fig. 2C), whereas Asn340 is located on the center of the outer β -sheet in strand I.

On the very top, both knobs carry a sialic acid binding site as identified by soaking the crystals with sialyl-lactose. The sialic acid and adjacent galactose residues are visible. This site is located in the highly positively charged surface at the top of the

FIG. 2. Structure of the Ad37 fiber head. (A) Ad37 knob trimer seen from the top. The positions of the bound sialic acid groups (green) and the residues differing in Ad37 (cyan) versus Ad19p (white) are shown in ball-and-stick representation. Circle highlights strands C' and I', which are typical of Ad37. (B) Representation of the CAR binding interface from the Ad12-CAR complex, with the backbone of the Ad12 knob shown in green and CAR shown in magenta. Contact residues of Ad12 are shown as ball-and-stick. The known structures of different serotypes have been aligned onto the Ad12 structure, and corresponding loops of Ad2 (yellow), Ad5 (blue), and Ad37 (red) are shown. EG' loops from another monomer are shown with brighter colors. (C) Stereogram of the Ad37 backbone (black). The trimer axis is tilted downward by 30° . The bound sialic acid residue, Lys240, and Asn340 are shown in ball-and-stick representations. The $C\alpha$ atoms of the residues corresponding to the CAR binding site of Ad12 (10) are shown as green spheres. (D) A subunit of Ad5 (blue) has been aligned with the corresponding $C\alpha$ atoms of Ad37 (black; corresponding to thick line in Fig. 2C). (E) A subunit of Ad3 (red) has been aligned with the corresponding $C\alpha$ atoms of Ad37 (black).

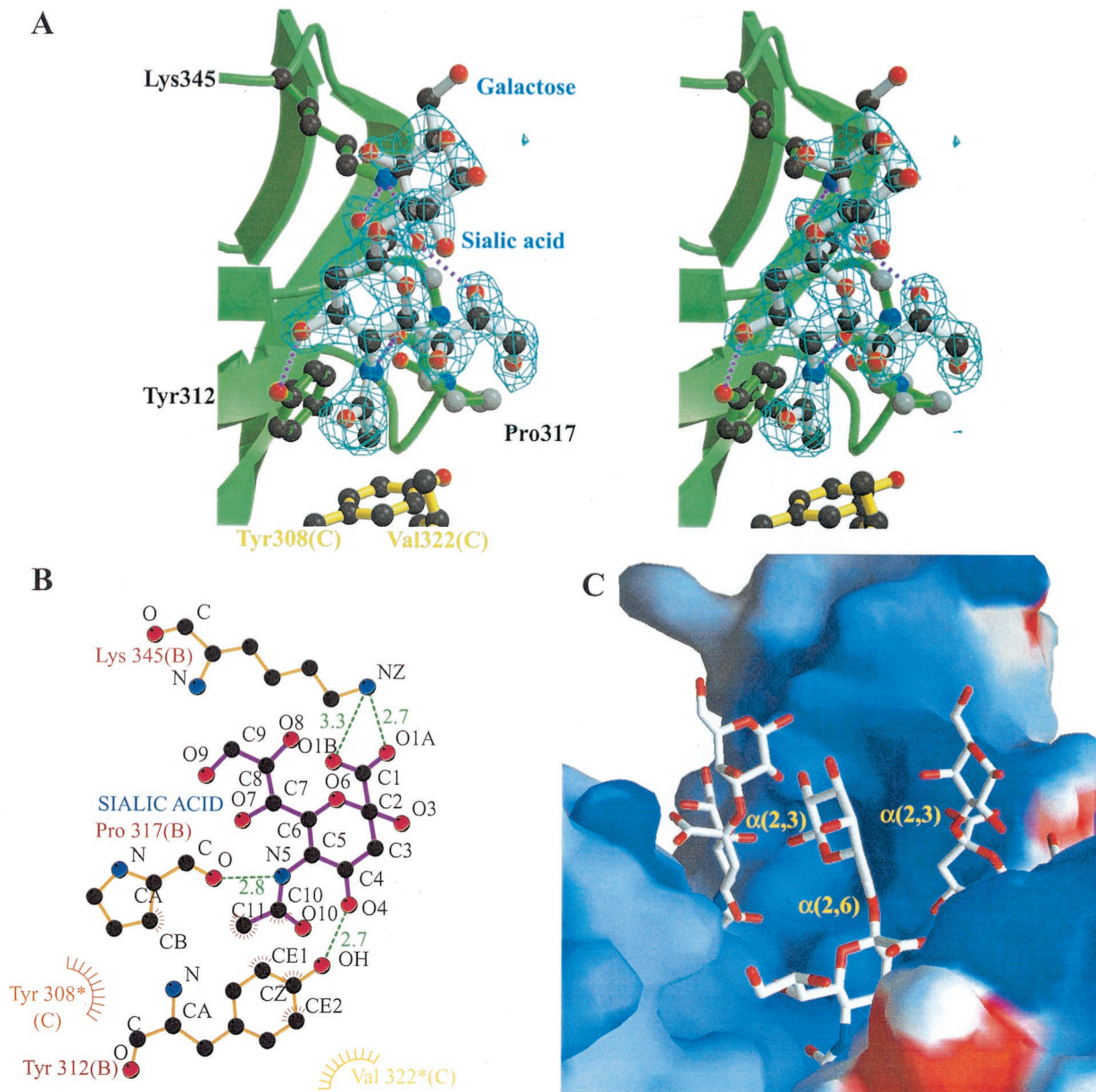


FIG. 4. Binding of sialyl-lactose. (A) Stereogram of bound sialyl- $\alpha(2,3)$ -lactose. In an omit map contoured at 3σ , the electron densities for the sialic acid and the galactose residues are shown. Purple dotted lines, hydrogen bonds. Residues from a neighboring subunit are shown in yellow. (B) Schema of the interactions between sialic acid (bonds in blue) and the fiber knob. The diagram was made with LIGPLOT (56). The loose hydrophobic interaction with Val322* (neighboring monomer) has been added by hand. Chain identifiers are given in parentheses. Hydrogen bonds are shown in green. Short strokes surrounding atoms or residue designations indicate hydrophobic contacts. (C) View onto the three sialic acid and galactose residues of sialyl-lactose located around the 3-fold axis. The galactose residues point away from the fiber knob, allowing the binding of sialic acid independently of its context.

trimer. The very basic surface of the molecule (Fig. 3) is characteristic of Ad37, compared, for example, with the structurally closely related surface of Ad5 fiber heads. Due to the two amino acid changes in the knob domain for Ad19p, the surface of this knob is less basic, even though the central part comprising the sialic acid binding sites remains unchanged.

The sialic acid residue of the sialyl-lactose used in the soaking experiment is in its usual chair conformation, presenting an axial carboxylic acid group. This group is recognized by hydrogen bonds to the amino group of Lys345 (Fig. 4A and B). The

hydroxyl group of Tyr312 hydrogen bonds to the hydroxyl group in position 4 of the sugar ring. The sugar ring is located on top of the stretch of residues from Pro317 to Pro320, where it excludes 2 to 3 water molecules and forms a sole hydrogen bond between its amide nitrogen and the carbonyl of Pro317. Tyr312, Tyr308 of another monomer, and to a lesser extent Pro317 and Val322 form a hydrophobic pocket, which binds the *N*-acetyl group. Interestingly, the glycerol moiety of the sialic acid is not involved in any direct interactions with the protein.

The compound used in the experiment is a mixture of $\alpha(2,3)$ - and $\alpha(2,6)$ -sialyl-lactose (about 85 and 15%, respectively). Surprisingly, for two of the binding sites on the trimer, the presence of the dominant $\alpha(2,3)$ -sialyl-lactose is observed, whereas for the third binding site, $\alpha(2,6)$ -sialyl-lactose is visible (Fig. 4C). This difference is due to an interaction of the galactose moiety with a symmetry-related protein for one of the binding sites, which creates the preference for $\alpha(2,6)$ -sialyl-lactose combined with a putative clash for an $\alpha(2,3)$ -linked lactose group. For the other molecules, the lactose moiety does not interact with the protein, but still the energy minimum of its glycosidic torsion angles, as discussed by Imberty et al. (29), fixes the orientation of the galactose residue. We observe $\phi = 56^\circ$, $\psi = 110^\circ$ and $\phi = 65^\circ$, $\psi = 111^\circ$ for the two $\alpha(2,3)$ -linked residues. Probably these two sites are also partially occupied with $\alpha(2,6)$ -sialyl-lactose, which is invisible due to low occupancy and the higher conformational flexibility of the $\alpha(2,6)$ -linkage.

The orientation of the carbohydrate molecule bound to the knob with the lactose groups pointing upward is as expected for a protein binding to terminal sialic acid (Fig. 4C). This indicates that interactions with other units of a complex carbohydrate structure are less likely.

There are no conformational changes upon sialic acid binding. The increase in resolution of crystals soaked with sialyl-lactose is probably due to improved crystal contact by the binding of the $\alpha(2,6)$ -sialyl-lactose in between two trimers.

Because a role of calcium in ligand binding had been shown by Wu et al. (59), we also collected data on crystals soaked simultaneously with 5 mM CaCl_2 and 10 mM sialyl-lactose. No extra electron density was present in this structure, excluding a role of calcium ions in sialic acid binding or in the fiber knob structure. The role of calcium must thus be limited to the stabilization of a cellular ligand or of its interaction with the fiber head.

The occupancy of the sialic acid binding site as a function of the ligand concentration was investigated in order to derive an estimate for the K_d of the sialyl-lactose–Ad37 knob complex. Analysis of the ligand occupancy is complicated due to the strong correlation between temperature factor and occupancy during refinement, even at resolutions between 2.2 and 1.5 Å. Furthermore, in the case of the free receptor, water molecules are bound in the ligand binding site, contributing to the electron density present at the position of the ligand. A careful treatment as described in Materials and Methods is thus necessary in order to obtain estimates for the real occupancy and the dissociation constant. Some limitations of the method still persist: the values may be affected by changes in the electrostatic potential due to neighboring molecules in the crystal, and the buffer used for crystallization may affect the dissociation constant. Still, the values obtained should give an idea of the order of magnitude of K_d . We obtain a K_d of 5 mM for Ad37 and a K_d of 7 mM for the Ad19p–sialyl-lactose complex (Fig. 5).

DISCUSSION

The two key residues of the sialic acid binding site, Tyr312 and Lys345, are conserved within the known fiber sequences of species D (Fig. 1). The other residues delimiting the binding site, Tyr308, Pro317, and Val322, are also largely conserved.

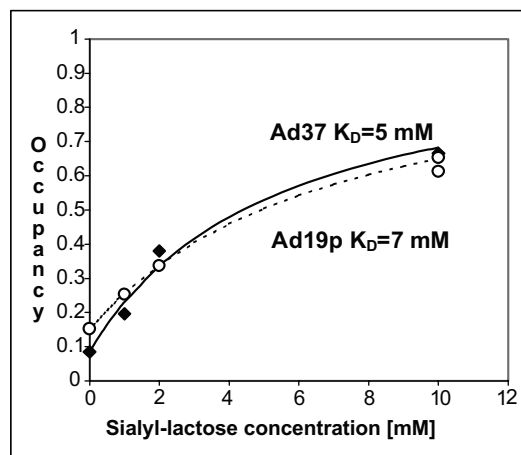


FIG. 5. Estimation of affinity for sialyl-lactose. The average occupancy of the bound $\alpha(2,3)$ -linked sialic acid groups was determined from the electron density in omit maps obtained after soaks with different concentrations of sialyl-lactose. Theoretical curves (dotted line, Ad19p; solid line, Ad37) and corresponding values for K_d are given. The formula and the method of occupancy calculation are given in Materials and Methods.

Pro317 can be replaced by alanine or isoleucine, but because the contact involves only the main chain carbonyl, the conformation of the main chain around this residue may still be the same. Most of the variation is found in the Ad15 sequence, with a replacement of Tyr308 by valine and of Pro317 by isoleucine. It is possible that this strain does not bind sialic acid, but the other strains in species D are very likely to do so.

The first viral sialic acid binding protein whose structure was solved was influenza virus hemagglutinin (58), followed (not considering neuraminidases) by polyomavirus VP1 protein (49) and the rotavirus VP4 sialic acid binding domain (20). These receptors differ greatly in their interactions with sialic acid. In particular, Ad37 fiber knobs and VP1 share the absence of interactions with the glycerol group. Influenza virus hemagglutinin and rotavirus VP4 form hydrogen bonds to the carboxyl group of sialic acid by means of the hydroxyl group of serine and a main-chain amide atom, whereas in VP1 and Ad37, fiber knob, arginine, and lysine residues form a salt bridge with the carboxyl group.

Recently, the structure of reovirus attachment protein $\sigma 1$ has been solved (13). This structure is related to the adenovirus fiber protein; it shows a similar organization in a trimeric head domain, formed here by three β -barrel structures with a topology similar to that of the adenovirus fiber knob, and a shaft domain formed, again, from a β -repeat similar to that observed for adenovirus (53). Unlike the Ad37 fiber head, this protein binds sialic acid through residues of its shaft region (14).

The K_d values of sialyl-lactose that we obtained, 5 mM for Ad37 and 7 mM for Ad19p, are not significantly different for the two strains and are in agreement with the 50% inhibitory concentration of 5 mM for Ad37 and sialyl-lactose determined in virus attachment inhibition assays (3). These values are in the range of dissociation constants observed for other sialic acid binding proteins. Influenza virus hemagglutinin binds sialyl-lactose with a K_d around 3 mM (strain X-31 [46]). In the case of the rotavirus VP8 core protein (21), a K_d around 1.2

mM has been obtained for various sialic acid-containing glycosides.

The conservation of the sialic acid binding site, the orientation of the bound sugar, and the order of magnitude of the affinity suggest that terminal sialic acid residues are certainly important in initial cell attachment of species D viruses. The interaction could be similar to the interaction of influenza virus with host cells, where sialic acid is used exclusively despite an affinity of the individual binding sites in the millimolar range.

Despite the presence of conserved amino acids in the sialic acid-interacting site of species D adenoviruses, other types besides Ad8, Ad19a, and Ad37 are rarely or never associated with EKC. Unlike other adenoviruses, these three viruses display the highest theoretical pI's in the knob domain (pI = 9.2). The pK_a of sialic acid is low (2.6), which strongly suggests the involvement of a charge-dependent interaction. In support of this suggestion, the binding of Ad37 to sialic acid has been shown to be sensitive to salt and negatively charged compounds (2).

Different possible explanations for high-affinity binding of virions with ocular tropism are as follows: (i) simply increased avidity due to multiple interactions between virions and cells, (ii) the additional positive charge of the fiber as already mentioned, (iii) the recognition *in vivo* of a modified, more complex sialylated carbohydrate structure, (iv) the simultaneous recognition of sialic acid and a protein surface, and (v) a second interaction with a protein molecule independent of sialic acid binding.

The difference in structure between serotype Ad37, with ocular tropism, and serotype Ad19p, without ocular tropism, is subtle and localized to the residues that are different in the two strains (Lys240Glu and Asn340Asp). The Lys240Glu mutation must be the more important, since it is sufficient to abolish the binding of Ad37 to Chang C cells (28).

Depending on the model, this mutation may affect binding in different ways. For example, Ad37 knobs could preferentially bind a modified sialic acid, such as $\alpha(2,8)$ -linked polysialic acid, which is possible because the glycerol chain of sialic acid does not contact the knob (Fig. 4). Such an interaction could be affected by the Lys240Glu mutation.

However, we cannot exclude the possibility that other, post-binding factors are involved in the determination of viral tropism. For example, events such as altered mRNA splicing (1) have been reported to abort adenovirus infections in monkey cells, and similar events may explain the limited ocular tropism exhibited by Ad19p as well.

The high similarity between the fiber knob of Ad37 and the structure of the CAR binding serotype Ad5 suggests that Ad37, as well as other subgroup D fiber heads, may be able to bind CAR, as has been shown by Kirby et al. (31) for Ad9. The CAR binding site is characterized by very poor conservation on the sequence level. The residues that have been identified as important for interaction with CAR (Fig. 1), either from the Ad12-CAR complex crystal structure (10) or from site-directed mutagenesis (42), are conserved in species D to the same extent as within the CAR binding serotypes (22). The only notable change concerns the conserved lysine at residue 301 in species D, which replaces the glutamic acid or threonine residue found in the CAR binding serotypes Ad12, Ad5, and Ad2. In the Ad37 fiber knob structure, the conformation of the

parts of the AB, CD, DE, and GE loops that are in contact with CAR in the Ad12-CAR complex is extremely similar to that of the CAR binding serotype Ad5 (Fig. 2B and D). The only difference concerns the conformation of the EG loop contributed by a neighboring monomer (Fig. 2B), but this loop may be flexible, leading to an induced fit contact with CAR.

Kirby et al. (31) showed the importance of the residue at position 227 (Ad37 numbering) for the interaction with CAR. A lysine in this position, such as is present in most viruses of the CAR binding species A and C, is very favorable for the interaction, whereas a glutamic acid residue such as is present in Ad9 is unfavorable. The species D sequences differ strongly in this position. None of the EKC-causing adenoviruses have this lysine, which may indicate that these viruses have evolved toward usage of a different receptor and a relatively weak affinity for CAR. This reduced affinity may still be sufficient for the interaction of the fiber with CAR in order to disrupt the epithelium, as has been shown by Walters et al. (57).

The CAR binding site, located on the side of the knob, and the sialic acid binding site, located on the very top of the knob, are distinct and do not overlap (Fig. 2C). Binding of these two ligands should be independent. It is unlikely that the carbohydrate structures of the two glycosylation sites on CAR can interact with the sialic acid binding site.

Recently, CD46 has been shown to be a ligand for Ad37 (59). The same protein has already been identified as the receptor for a number of species B viruses (24) such as Ad11p (47) and Ad3 (48). Alignment of the amino acid sequences of the knob of Ad37 with species B representatives (Ad3 and Ad11) shows relatively little identity (28 and 29%, respectively, compared to 48% for Ad5 [Fig. 1]). Furthermore, alignment of the α backbone with Ad3, the only known 3-dimensional structure of species B, shows little conservation in the loop regions (Fig. 2E). The Ad3 structure is also unique due to the hydrophobic character of its surface, which is more or less conserved within species B (22), a feature not shared by Ad37 (data not shown). It is therefore rather unlikely that there is a shared CD46 binding site for the two virus species.

The two N-terminal repeats of CD46 with a known X-ray structure (12,) to which measles virus binds, have all the features expected for a suitable ligand for Ad37 fibers: a negative surface charge complementary to the positive charge of the fiber head (by analysis with GRASP [40] [data not shown]), a potentially trimeric structure (12), and sialic acid-containing complex carbohydrates. However, CD46 is expressed on all nucleated cells, and binding of Ad37 to CD46 cannot alone explain the specific ocular tropism. Moreover, CD46 is present in ocular fluid, indicating that it may inhibit rather than promote Ad37 binding (16).

The Ad37 knob structure will be valuable for the design of antiadenovirus drugs, in particular drugs directed against EKC which block the sialic acid binding site or another interaction with a cellular ligand. However, only a structure of the complex with the entire cellular receptor molecule will yield a definitive answer about the mode of interaction.

ACKNOWLEDGMENTS

We thank the JSBG for beamtime, Antoine Royant for assistance on beamline ID29, and G. Kleywegt for the suggestion on how to integrate electron densities.

This work has been supported by The Swedish Research Council (grant 521-2002-5981) and Magnus Bergwalls Stiftelse.

REFERENCES

- Anderson, K. P., and D. F. Klessig. 1984. Altered mRNA splicing in monkey cells abortively infected with human adenovirus may be responsible for inefficient synthesis of the virion fiber polypeptide. *Proc. Natl. Acad. Sci. USA* **81**:4023–4027.
- Arnberg, N., A. H. Kidd, K. Edlund, J. Nilsson, P. Pring-Akerblom, and G. Wadell. 2002. Adenovirus type 37 binds to cell surface sialic acid through a charge-dependent interaction. *Virology* **302**:33–43.
- Arnberg, N., P. Pring-Akerblom, and G. Wadell. 2002. Adenovirus type 37 uses sialic acid as a cellular receptor on Chang C cells. *J. Virol.* **76**:8834–8841.
- Arnberg, N., A. H. Kidd, K. Edlund, F. Olfat, and G. Wadell. 2000. Initial interactions of subgenus D adenoviruses with A549 cellular receptors: sialic acid versus α v integrins. *J. Virol.* **74**:7691–7693.
- Arnberg, N., K. Edlund, A. H. Kidd, and G. Wadell. 2000. Adenovirus type 37 uses sialic acid as a cellular receptor. *J. Virol.* **74**:42–48.
- Arnberg, N., Y. Mei, and G. Wadell. 1997. Fiber genes of adenoviruses with tropism for the eye and the genital tract. *Virology* **6**:239–244.
- Bailey, A., and V. Mautner. 1994. Phylogenetic relationships among adenovirus serotypes. *Virology* **205**:438–452.
- Benkő, M., B. Harrach, and W. C. Russell. 2000. Family *Adenoviridae*, p. 227–238. *In* M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wickner (ed.), *Virus taxonomy: classification and nomenclature of viruses*. Seventh report of the International Committee on Taxonomy of Viruses. Academic Press, Inc., San Diego, Calif.
- Bergelson, J. M., J. A. Cunningham, G. Droguett, E. A. Kurt-Jones, A. Krithivas, J. S. Hong, M. S. Horwitz, R. L. Crowell, and R. W. Finberg. 1997. Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science* **275**:1320–1323.
- Bewley, M. C., K. Springer, Y. B. Zhang, P. Freimuth, and J. M. Flanagan. 1999. Structural analysis of the mechanism of adenovirus binding to its human cellular receptor, CAR. *Science* **286**:1579–1583.
- Brünger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren. 1994. Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**:905–921.
- Casasnovas, J. M., M. Larvie, and T. Stehle. 1999. Crystal structure of two CD46 domains reveals an extended measles virus-binding surface. *EMBO J.* **18**:2911–2922.
- Chappell, J. D., A. E. Prota, T. S. Dermody, and T. Stehle. 2002. Crystal structure of reovirus attachment protein σ 1 reveals evolutionary relationship to adenovirus fiber. *EMBO J.* **21**:1–11.
- Chappell, J. D., J. L. Duong, B. W. Wright, and T. S. Dermody. 2000. Identification of carbohydrate-binding domains in the attachment proteins of type 1 and type 3 reoviruses. *J. Virol.* **74**:8472–8479.
- Chroboczek, J., R. Ruigrok, and S. Cusack. 1995. Adenovirus fiber. *Curr. Top. Microbiol. Immunol.* **199**:163–200.
- Cocuzzi, E., L. B. Szczołka, W. G. Brodbeck, D. S. Bardenstein, T. Wei, and M. E. Medof. 2001. Tears contain the complement regulator CD59 as well as decay-accelerating factor (DAF). *Clin. Exp. Immunol.* **123**:188–195.
- Cohen, C. J., J. T. Shieh, R. J. Pickles, T. Okegawa, J. T. Hsieh, and J. M. Bergelson. 2001. The coxsackievirus and adenovirus receptor is a transmembrane component of the tight junction. *Proc. Natl. Acad. Sci. USA* **98**:15191–15196.
- Collaborative Computational Project, Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**:760–763.
- Dörig, R. E., A. Marciel, A. Chopra, and C. D. Richardson. 1993. The human CD46 molecule is a receptor for measles virus (Edmonston strain). *Cell* **75**:295–305.
- Dormitzer, P. R., Z. Y. Sun, G. Wagner, and S. C. Harrison. 2002. The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site. *EMBO J.* **21**:885–897.
- Dormitzer, P. R., Z. Y. Sun, O. Blixt, J. C. Paulson, G. Wagner, and S. C. Harrison. 2002. Specificity and affinity of sialic acid binding by the rhesus rotavirus VP8* core. *J. Virol.* **76**:10512–10517.
- Durmort, C., C. Stehlin, G. Schoehn, A. Mitraki, E. Drouet, S. Cusack, and W. P. Burmeister. 2001. Structure of the fiber head of Ad3, a non-CAR-binding serotype of adenovirus. *Virology* **285**:302–312.
- Ford, E., K. E. Nelson, and D. Warren. 1987. Epidemiology of epidemic keratoconjunctivitis. *Epidemiol. Rev.* **9**:244–261.
- Gaggar, A., D. M. Shayakhmetov, and A. Lieber. 2003. CD46 is a cellular receptor for group B adenoviruses. *Nat. Med.* **9**:1408–1412.
- Gordon, Y. J., K. Aoki, and P. R. Kitchington. 1996. Adenovirus keratoconjunctivitis, p. 877–894. *In* J. S. Pepose, G. N. Holland, and K. R. Wilhelmus (ed.), *Ocular infection and immunity*. Mosby, St. Louis, Mo.
- Gouet, P., E. Courcelle, D. I. Stuart, and F. Metz. 1999. ESPript: multiple sequence alignments in PostScript. *Bioinformatics* **15**:305–308.
- Howitt, J., M. C. Bewley, V. Graziano, J. M. Flanagan, and P. Freimuth. 2003. Structural basis for variation in adenovirus affinity for the cellular coxsackievirus and adenovirus receptor. *J. Biol. Chem.* **278**:26208–26215.
- Huang, S., V. Reddy, N. Dasgupta, and G. R. Nemerow. 1999. A single amino acid in the adenovirus type 37 fiber confers binding to human conjunctival cells. *J. Virol.* **73**:2798–2802.
- Imberty, A., C. Gautier, J. Lescar, S. Pérez, L. Wyns, and R. Loris. 2000. An unusual carbohydrate binding site revealed by the structures of two *Maackia amurensis* lectins complexed with sialic acid-containing oligosaccharides. *J. Biol. Chem.* **275**:17541–17548.
- Jones, T. A., J.-Y. Zou, S. W. Cowan, and M. Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and location of errors in these models. *Acta Crystallogr. A* **47**:110–119.
- Kirby, I., R. Lord, E. Davison, T. J. Wickham, P. W. Roelvink, I. Kovesdi, B. J. Sutton, and G. Santis. 2001. Adenovirus type 9 fiber knob binds to the coxsackie B virus-adenovirus receptor (CAR) with lower affinity than fiber knobs of other CAR-binding adenovirus serotypes. *J. Virol.* **75**:7210–7214.
- Kirby, I., E. Davison, A. J. Beavil, C. P. Soh, T. J. Wickham, P. W. Roelvink, I. Kovesdi, B. J. Sutton, and G. Santis. 2000. Identification of contact residues and definition of the CAR-binding site of adenovirus type 5 fiber protein. *J. Virol.* **74**:2804–2813.
- Kirby, I., E. Davison, A. J. Beavil, C. P. Soh, T. J. Wickham, P. W. Roelvink, I. Kovesdi, B. J. Sutton, and G. Santis. 1999. Mutations in the DG loop of adenovirus type 5 fiber knob protein abolish high-affinity binding to its cellular receptor CAR. *J. Virol.* **73**:9508–9514.
- Kleywegt, G. J., and T. A. Jones. 1999. Software for handling macromolecular envelopes. *Acta Crystallogr. D* **55**:941–944.
- Kleywegt, G. J., and T. A. Jones. 1996. xdlMAPMAN and xdlDATAMAN: programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. *Acta Crystallogr. D* **52**:826–828.
- Leslie, A. G. W. 1999. Integration of macromolecular diffraction data. *Acta Crystallogr. D* **55**:1696–1702.
- Mei, Y. F., and G. Wadell. 1995. Molecular determinants of adenovirus tropism. *Curr. Top. Microbiol. Immunol.* **199**:213–228.
- Naniche, D., G. Varior-Krishnan, F. Cervoni, T. F. Wild, B. Rossi, C. Raubourdin-Combe, and D. Gerlier. 1993. Human membrane cofactor protein (CD46) acts as a cellular receptor for measles virus. *J. Virol.* **67**:6025–6032.
- Navaza, J. 1994. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**:157–163.
- Nicholls, A., K. Sharp, and B. Honig. 1991. Protein folding and association: insight from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**:281–296.
- Perrakis, A., R. Morris, and V. Lamzin. 1999. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6**:458–463.
- Roelvink, P. W., G. Mi Lee, D. A. Einfeld, I. Kovesdi, and T. J. Wickham. 1999. Identification of a conserved receptor-binding site on the fiber proteins of CAR-recognizing adenoviridae. *Science* **286**:1568–1571.
- Roelvink, P. W., A. Lizonova, J. G. Lee, Y. Li, J. M. Bergelson, R. W. Finberg, D. E. Brough, I. Kovesdi, and T. J. Wickham. 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**:7909–7915.
- Roizman, B. 1996. *Adenoviridae*, p. 2111–2171. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), *Fields virology*, 3rd ed., vol. 2. Lippincott-Raven, Philadelphia, Pa.
- Santoro, F., P. E. Kennedy, G. Locatelli, M. S. Malnati, E. A. Berger, and P. Lusso. 1999. CD46 is a cellular receptor for human herpesvirus 6. *Cell* **99**:817–827.
- Sauter, N. K., M. D. Bednarski, B. A. Wurzburg, J. E. Hanson, G. M. Whitesides, J. J. Skehel, and D. C. Wiley. 1989. Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study. *Biochemistry* **28**:8388–8396.
- Seegerman, A., J. P. Atkinson, M. Marttila, V. Dennerquist, G. Wadell, and N. Arnberg. 2003. Adenovirus type 11 uses CD46 as a cellular receptor. *J. Virol.* **77**:9183–9191.
- Sirena, D., B. Lilienfeld, M. Eisenhut, S. Kälin, K. Boucke, R. R. Beerli, L. Vogt, C. Ruedl, M. F. Bachmann, U. F. Greber, and S. Hemmi. 2004. The human membrane cofactor CD46 is a receptor for the species B adenovirus serotype 3. *J. Virol.* **78**:4454–4462.
- Stehle, T., and S. C. Harrison. 1997. High-resolution structure of a polyomavirus VP-1 oligosaccharide complex: implications for assembly and receptor binding. *EMBO J.* **16**:5139–5148.
- Stevens, R. C. 2000. Design of high-throughput methods of protein production for structural biology. *Struct. Fold. Des.* **8**:177–185.
- Tomko, R. P., R. Xu, and L. Philipson. 1997. HCAR and MCAR: the human and mouse cellular receptors for subgroup C adenoviruses and group B coxsackieviruses. *Proc. Natl. Acad. Sci. USA* **94**:3352–3356.
- van Raaij, M. J., E. Choiun, H. van der Zandt, J. M. Bergelson, and S.

- Cusack. 2000. Dimeric structure of the coxsackie and adenovirus receptor D1 domain at 1.7 Å resolution. *Struct. Fold. Des.* **8**:1147–1155.
53. van Raaij, M. J., A. Mitraki, G. Lavigne, and S. Cusack. 1999. A triple beta-spiral in the adenovirus fiber shaft reveals a new structural motif for a fibrous protein. *Nature* **401**:935–938.
54. van Raaij, M. J., N. Louis, J. Chroboczek, and S. Cusack. 1999. Structure of the human adenovirus serotype 2 fiber head domain at 1.5 Å resolution. *Virology* **262**:333–343.
55. Wadell, G., M. L. Hammarskjold, G. Winberg, T. M. Varsanyi, and G. Sundell. 1980. Genetic variability of adenoviruses. *Ann. N. Y. Acad. Sci.* **354**:16–42.
56. Wallace, A. C., R. A. Laskowski, and J. M. Thornton. 1995. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng.* **8**:127–134.
57. Walters, R. W., P. Freimuth, T. O. Moninger, I. Ganske, J. Zabner, and M. J. Welsh. 2002. Adenovirus fiber disrupts CAR-mediated intercellular adhesion allowing virus escape. *Cell* **110**:789–799.
58. Weis, W., J. H. Brown, S. Cusack, J. C. Paulson, J. J. Skehel, and D. C. Wiley. 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature* **333**:426–431.
59. Wu, E., S. A. Trauger, L. Pache, T.-M. Mullen, D. J. Von Seggern, G. Siuzdak, and G. R. Nemerow. 2004. Membrane cofactor protein is a receptor for adenoviruses associated with epidemic keratoconjunctivitis. *J. Virol.* **78**:3897–3905.
60. Wu, E., J. Fernandez, S. K. Fleck, D. J. Von Seggern, S. Huang, and G. R. Nemerow. 2001. A 50-kDa membrane protein mediates sialic acid-independent binding and infection of conjunctival cells by adenovirus type 37. *Virology* **279**:78–89.
61. Xia, D., L. J. Henry, R. D. Gerard, and J. Deisenhofer. 1994. Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution. *Structure* **2**:1259–1270.