Adenovirus serotype 26 utilises sialic acid bearing glycans as a primary cell entry receptor

Alexander T. Baker[1], Rosie Mundy[1], James Davies[1], Pierre J. Rizkallah[2], Alan L Parker[1].

[1] Division of Cancer and Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff, CF14 4XN, UK.

[2] Division of Infection and Immunity, School of Medicine, Cardiff University, Heath Park, Cardiff, CF14 4XN, UK.

Correspondence:
Dr Alan Parker
Institute of Cancer and Genetics
Cardiff University
Heath Park
Cardiff
CF14 4XN
Telephone: +44 2922 510 231
Email: ParkerAL@cardiff.ac.uk

Running title: Ad26 binds sialic acid for cell entry.

Keywords: Adenovirus, sialic acid, receptor, crystal structure

Abstract word count: 150

Main text word count: 5009

Figure legend word count: 462

Supplemental figure legend word count: 404

Figure count: 5 + 4 supplementary + 1 supplementary table

Abstract:

Adenoviruses are clinically important agents. They cause respiratory distress, gastroenteritis, and epidemic keratoconjunctivitis (EKC). As non-enveloped, double stranded DNA viruses, they are easily manipulated, making them popular vectors for therapeutic applications, including vaccines. Species D adenovirus serotype 26 (HAdV-D26) is both a cause of EKC and other disease, and a promising vaccine vector. HAdV-D26 derived vaccines are under investigation as protective platforms against HIV, Zika, RSV infections and are in Phase-III clinical trials for Ebola.

We recently demonstrated that HAdV-D26 does not utilise CD46 or desmoglein 2 as entry receptors, whilst the putative interaction with Coxsackie and Adenovirus Receptor (CAR) is low affinity and unlikely to represent the primary cell receptor.

Here, we definitively establish sialic acid as the primary entry receptor utilised by HAdV-D26. We demonstrate removal of cell surface sialic acid inhibits HAdV-D26 infection and provide a high-resolution crystal structure of HAdV-D26 fiber-knob in complex with sialic acid.

Introduction:

Adenoviruses are clinically significant, both as human pathogens, and as platforms for therapeutic applications. As pathogens, human adenoviruses (HAdV) have been isolated from severe infections in both immunocompromised and healthy individuals(*1*, *2*). Some adenoviruses have been associated with acute infection of the eye(*3*), respiratory(*4*, *5*), and gastrointestinal tract(*1*, *6*). In rare cases infections prove fatal, as observed in a recent neonatal infection of HAdV-D56(*7*), among adult patients in New Jersey with HAdV-B7d(*8*), and infamously with HAdV-E4 where large epidemics of adenovirus infection have been seen in military recruits(*9*, *10*). However, most fatal infections are observed in immunocompromised individuals(*11*), such as those suffering from graft versus host disease (GVHD)(*12*) or HIV(*13*, *14*).

Adenoviruses are classified into 7 species (A-G)(*15*), and between 57 and 90 serotypes depending on the taxonomic definitions used(*16*, *17*). Some adenoviruses have been studied in detail, having well defined receptor tropisms, including as coxsackie and adenovirus receptor (CAR)(*18*, *19*), CD46 (Major Complement Protein, MCP)(*20–22*), desmoglein 2 (DSG2)(*23–25*), or sialic acid bearing glycans(*26–29*). However, most serotypes have low seroprevalence in the population(*30–33*), though this is varies significantly by geographical location(*34*, *35*). Their rarity means many serotypes remain understudied, with poorly defined primary receptor interactions. This is especially true of the species D adenoviruses (HAdV-D); the largest of the adenoviral species, containing 35 of the 57 canonical serotypes(*17*).

Species D adenoviruses are associated with several pathogenicities. HAdV-D56 is a potentially fatal emergent respiratory pathogen comprised of a recombination between 4 species D adenoviruses(*7*). Opportunistic adenovirus infection isolated from HIV/AIDS patients are most commonly from species D, where they are associated with prolonged shedding in the gastrointestinal tract(*13*). HAdV-D has also been associated with genital disease(*36*, *37*). The species D adenoviruses are best known, however, for causing epidemic keroconjunctivitis (EKC) infections, which is endemic, but not isolated, to Japan(*38*, *39*). Classically the primary EKC causing adenoviruses have been HAdV-D8(*40–42*), 37(*42–44*), and 64(*42*, *45*) (previously classified as 19a(*46*)). More recently other species D adenovirus have been associated with EKC, including HAdV-D53(previously classified as HAdV-D22/H8)(*47*, *48*), 54(*49*), 56(*50*),

Their double stranded DNA genome makes them readily amenable to genetic modification(*51*), and therefore has made them attractive candidates for genetic manipulation for therapeutic applications in cancer (oncolytic viruses)(*52*) and as vaccine vectors(*53*, *54*). Species D adenoviruses are of special interest as vaccine vectors. Their ability to induce robust cellular and humoral immunogenic responses

in humans, coupled with low seroprevalence rates in the general population(*30*, *33*) makes them attractive platforms for vaccines, as evidenced by their progression through clinical trials for HIV(*55*, *56*), Zika(*57*), and Ebola treatment(*58*, *59*). However, there remains a lack of understanding regarding their basic biology and mechanisms of cellular infection. This is exemplified by Adenovirus serotype 26 (HAdV-D26), which is being investigated as a vaccine vector for zika(*57*), HIV(*60*), respiratory syncytial virus(*61*), and has entered phase III clinical trials as an Ebola vaccine(*58*).

Despite its clinical success, recent findings further highlight the lack of clarity over the primary receptor usage of HAdV-D26. It is now clear that, despite previous publications to the contrary, HAdV-D26 cannot engage CD46 as a primary cellular entry receptor(*62*). Instead, the HAdV-D26 fiber-knob protein (HAdV-D26K) may engage CAR as a primary receptor, although the affinity of this interaction is attenuated compared to the classical HAdV-C5 interaction with CAR due to the presence of an extended HAdV-D26 fiber-knob DG loop, which sterically inhibits the interaction with CAR(*62*). The deduced low affinity of the interaction between CAR and HAdV-D26 fiber-knob make it unlikely that CAR represents the definitive primary receptor of HAdV-D26.

Here, we conclusively demonstrate that HAdV-D26 utilises sialic acid bearing glycans as a primary entry receptor, and that this interaction can form a productive infection. We deduce the structure of HAdV-D26K in complex with sialic acid (Neu5Ac), demonstrating a similar topology to the known sialic acid interacting adenovirus HAdV-D37 fiber-knob in the sialic acid binding pocket, but highlight crucial mechanistic differences likely to enhance HAdV-D26 affinity for sialic acid compared to other serotypes.

Results:

HAdV-D26K has an electrostatic profile permissive to sialic acid interaction:

Our recent findings rule out any role for DSG-2 or CD46 in HAdV-D26 infection, whilst the low affinity of the interaction between HAdV-D26K and CAR made it an unlikely primary cell entry receptor. Previous amino acid sequence alignments demonstrated little conservation of sialic acid binding residues with the fiber-knob domains of the known sialic acid utilising adenoviruses HAdV-G52SFK (short fiber-knob) or Canine adenovirus serotype 2 (CAV-2)(*62*). However, these alignments indicated HAdV-D37, known to bind sialic acid in the apex of the fiber-knob, bore some similarity at a sequence level. We sought to evaluate the ability of HAdV-D26K to interact with the remaining previously described adenovirus receptor, sialic acid.

HAdV-D37 fiber-knob is identical to that of HAdV-D64, and highly homologous to HAdV-D8 (Fig.1A). These three viruses have been shown to cause epidemic keratoconjunctivitis (EKC) and to interact

with sialic acid. The closely related HAdV-D19p, differing from HAdV-D64 at only two residues, has also been shown to bind sialic acid, but does not cause EKC. We compared HAdV-D26K to these viruses to determine if a similar binding mechanism was possible.

These sialic acid binding viruses all have highly negative predicted isoelectric points (pI) (Fig.1A). We calculated the surface electrostatic potentials of these fiber-knob proteins, at pH7.35 to simulate the pH of extracellular fluid, using previously published crystal structures where available. There is no published structure of HAdV-D8K, so we generated a homology model based on the closest known relative with a crystal structure (Fig.1B).

The viruses are highly basic, with a concentration of positive charge in the central depression around the 3-fold axis corresponding to the previously reported sialic acid binding sites (Fig.1B-D). We observed that HAdV-D8 has the most basic surface potential (Fig.1B), followed by HAdV-D37/64 (Fig.1C). HAdV-D19p is less basic, due to the two amino acid substitutions, compared to HAdV-D37/64, though the central depression is unaffected, as has previously been noted (Fig.1D)(*63*).

HAdV-D26K has a lower predicted pI, 6.49, and less positive surface potential (Fig.1A,E). However, the central depression of HAdV-D26K remains basic around the region where sialic acid is observed to bind in HAdV-D19p and HAdV-D37. HAdV-D26 retains the charge needed for sialic acid binding in the apex of the protein in the context of an otherwise acidic protein (Fig.1F).

HAdV-D26 requires cell surface sialic acid for efficient infection:

Sequence alignment of HAdV-D26K with these known sialic acid utilising viruses, bearing a positively charged apex, showed conservation of key binding residues between serotypes (Fig.2A). We observe complete conservation of Tyr130, and Lys165 across the 4 serotypes, and conservation of Asp128 with HAdV-D8 (Fig.2A). Further, while Tyr135 is not conserved in HAdV-D26, inspection of the crystal structure of HAdV-D37K and HAdV-D19p in complex with sialic acid (PDB 1UXA and 1UXB, respectively)(*63*) reveals this to be a main chain oxygen contact, positioned similarly in HAdV-D26, and can be considered homologous.

To investigate the ability of HAdV-D26 to utilise sialic acid as a cell entry receptor we used a replication incompetent HAdV-C5 vector pseudotyped with the HAdV-D26 fiber-knob, expressing a GFP transgene. We performed infectivity studies in three cell lines, with and without pre-treatment with neuraminidase to remove cell surface sialic acid. The tested cell lines could be infected by the CD46 (Fig.2B) or CAR (Fig.2C) mediated pathways to some extent, by HAdV-C5/B35K or HAdV-C5, respectively. However, infection via these routes was uninhibited by neuraminidase treatment. Transduction efficiency of HAdV-C5 and HAV-C5/B35K was actually enhanced by neuraminidase

143 treatment in some cases; an effect which has been previously observed(*27*, *64*). This has been
144 suggested to be due to a reduction in the electrostatic repulsion of the negatively charged capsid of
145 HAdV-C5.

146 Infection by the HAdV-C5/D26K pseudotype was significantly reduced in all three cell lines following
147 treatment with neuraminidase (Fig.2D). This inhibition is significant ($P<0.005$), resulting in >5-fold
148 decrease in infection, in all three cell lines tested. These data indicate that HAdV-C5/D26K is likely to
149 be utilising the sialic acid mediated pathway for infection, not CD46 or CAR.

150 **HAdV-D26 forms a stable complex with sialic acid:**

151 We crystallised HAdV-D26K in complex with sialic acid in order to clarify the mechanism of interaction.
152 Refinement of structures generated from HAdV-D26K crystals soaked in sialic acid shows electron
153 density for a small molecule ligand in the apical depression (Fig.3A), this is best described by a racemic
154 mixture of α and β anomers, in conjunction with double conformations of sialic acid (Fig.3B). The cubic
155 space group (supplementary table 1) enabled assembly of the biological trimer. We observed three
156 copies of sialic acid bound within the apex of the fiber-knob trimer (Fig.3C), as previously observed in
157 HAdV-D37 and HAdV-D19p.

158 Sialic acid binding was observed in structures crystallised at both pH8.0 (PDB 6QU6) and pH4.0 (PDB
159 6QU8). Observation of sialic acid density at high σ-values suggests a highly stable interaction
160 (Supplementary Fig.1). Electron density demonstrates the C2 carboxyl and OH groups in two
161 conformations, and the C6 glycerol group is flexible, with the C7-C8 bond rotating to alter the
162 orientation of the glycerol arm relative to the pyranose ring and binding pocket (Fig.3A,B,
163 Supplementary Fig.1). The glycerol group exhibits further flexibility at the C8-C9 bond, making the
164 terminal oxygen mobile. The distribution of the density for the glycerol group is different at each pH
165 (Supplementary Fig.1), suggesting pH could affect the preferred mode of interaction.

166 The most biologically relevant sialic acid conformation places the carboxyl group axial to the chair-
167 conformation pyranose ring (Supplementary Fig.2), leaving the OH group pointing away from the fiber-
168 knob and free to form an α(2)-glycosidic bond as part of a glycan. This is suggestive of a terminal sialic
169 acid residue, as the chain can extend out of the central depression, as was observed in the previously
170 described HAdV-D37K:GD1a glycan structure(*63*, *65*).

171 **HAdV-D26 possesses a sophisticated sialic acid binding pocket:**

172 Comparison between the HAdV-D26K and HAdV-D37K, the best described of the sialic acid binding
173 adenoviruses, reveals several sialic acid contacts are conserved (Fig.4A,B). Lys349 and Tyr314 are
174 identical, and while Lys349 exhibits some flexibility, all observed lysine conformations form a contact

with the carboxyl-group of the sialic acid (Supplementary Fig.3). Whilst Thr319 is not conserved in HAdV-D37 (which has a proline at this position), the main chain oxygen contact to the N-Acetyl nitrogen is spatially similar, so the bond can be considered homologous.

The HAdV-D26K sialic acid interface forms further contacts with sialic acid that are not observed in HAdV-D37K (Fig.4A, B). HAdV-D26K contacts the N-Acetyl oxygen of sialic acid using Asn312, which forms a polar contact and a water-bridge (Fig.4A). The comparable residue in HAdV-D37K, Thr310, is too short to form a direct polar interaction (Fig.4B), instead utilising a pair of water-bridges.

In HAdV-D37 the glycerol arm of sialic acid was only contacted by a water-bridge between Ser344 and the C7-OH. However, in HAdV-D26 all 3 OH groups in the glycerol arm form contacts. C7-OH is coordinated by water-bridges to both Asn312, and Gln348. C8-OH forms a water-bridge with Thr319, and C9-OH forms both a water-bridge and a polar contact directly to Gln348. Like Thr310, the serine belonging to HAdV-D37 at position 344 is too short to form a polar bond equivalent to the one with Gln348.

Notably, the density for the glycerol arm of sialic acid suggests several possible conformations (supplementary fig.2) which can be interpreted as flexibility. However, we suggest that, in HAdV-D26, this is unlikely since it is so well coordinated in all conformations observed, at both pH8.0 and pH4.0 (supplementary fig.3). We propose that HAdV-D26K can form a stable interaction with the glycerol arm, regardless of the specific confirmation. The variable density can be explained as the average distribution (or partition) of the different discrete positions.

We also observe a hydrophobic interaction in HAdV-D26 with the N-Acetyl methyl group at C11 (Fig.4C). A similar hydrophobic interaction is seen in HAdV-D37, where Tyr312 and Val322 form a hydrophobic patch (Fig.4D), but the HAdV-D37 interaction appears to be more selective, where Ile310, and Ile324 form a hydrophobic cradle around the methyl group (Fig.4C).

HAdV-D26 binds sialic acid through an induced fit mechanism:

We observe split density for Gln348 in both pH8.0 (Fig.5A) and pH4.0 (Fig.5B). Whilst conformation A can form polar contacts with sialic acid, conformation B points into the solvent and cannot. It is possible that Gln348 is flexible, but then is attracted to the charged density of the glycerol arm upon sialic acid interaction. We also observe greater occupancy of conformation A in the pH8.0 structure (approximately 0.7) while at pH4.0 the occupancy is evenly split. This suggests that the interaction may be more stable at higher pH, such as that associated with the pH found at the cell surface.

Ile324, which is seen to be involved in hydrophobic interactions with the N-Acetyl methyl group (Fig.4C), can also have multiple conformations. In an unliganded structure of HAdV-D26 fiber-knob

(PDB 6FJO) the long arm of Ile324 is seen to rotate (Fig.5C). However, in the ligated structure Ile324 occupies a single conformation (Fig.5D) forming a cradle. This creates a larger hydrophobic patch and restricts the methyl group in space by pinching it between the pair of hydrophobic isoleucines, anchoring the N-Acetyl group.

Discussion:

Other adenoviruses have previously been shown to interact with sialic acid. These include CAV2(*66*), Turkey adenovirus 3 (TAdV-3)(*67*), and HAdV-G52 short fiber-knob(*26*, *68*), but these viruses interact with sialic acid in lateral regions of the fiber-knob, dissimilar from HAdV-D26K. Four other human adenoviruses fiber-knob proteins (HAdV-D8/19p/37/64K) have been previously shown to utilise sialic acid, binding in the apical region. These viruses have high sequence similarity to each other, but not to HAdV-D26K, though they all share key sialic acid contact residues (Fig.1A,2A).

The structure of HAdV-D8 has not been determined, either alone or in complex with sialic acid, but infection by HAdV-D8 is sensitive to neuraminidase treatment suggesting sialic acid utilisation(*69*). Furthermore, HAdV-D8K has very high sequence homology and shared sialic acid contact residues with HAdV-D19p/37K making it logical to expect a similar interaction mechanism. In support of this we observe a similar electrostatic profile in the modelled fiber-knob as seen in HAdV-D37/64 (Fig.1B,C). HAdV-D64 has an identical fiber-knob domain to that of HAdV-D37, so fiber-knob interactions with sialic acid are likely to be conserved between these serotypes. HAdV-D26K conserves the key region of positive potential in the apical depression, but in the context of an otherwise more acidic protein (Fig.1E).

Inspection of the sialic binding pocket of HAdV-D26K reveals a much more complex mechanism of interaction than that previously reported for HAdV-D37K (Fig.4)(*63*). The overall topology of the pocket is similar, with hydrophobic residues around the N-Acetyl group and polar contacts between the carboxyl and C4-OH group. However, HAdV-D26K has several differences which increase the number of contacts between the sialic acid and the fiber-knob.

Subtle sequence changes enable more numerous interactions between HAdV-D26K and sialic acid than are possible in HAdV-D37K. In HAdV-D37 Pro317 forms a main chain oxygen contact to the nitrogen of sialic acid, however it also creates tension which rotates the N-terminal residue away from the carboxyl group of sialic acid. In HAdV-D26K, Tyr320, which is C-terminal of the Thr319 that is equivalent to Pro317 in HAdV-D37K does not create this tension and enables the main chain oxygen at position 320 to contact the sialic acid carboxyl group. Thr319, also forms a water-bridge with the C8-OH group, helping to stabilise the glycerol side-chain (Fig.4A).

239 This is one of several examples of HAdV-D26K being better evolved to contact sialic acid. The
240 substitution of the Thr310 and Ser344 found in HAdV-D37K for longer charged residues (Asn312 and
241 Gln348, respectively) in HAdV-D26K enables direct polar contacts, as well as additional water-bridge
242 contacts. Substitution of Tyr308 and Val322 for more hydrophobic isoleucine residues in HAdV-D26
243 (Ile 310 and Ile324, respectively) creates a hydrophobic indentation better tailored to fit around the
244 N-Acetyl methyl group.

245 The high resolution of the datasets generated to determine the sialic acid bound HAdV-D26K structure
246 enables visualisation of multiple residue conformations with partial occupancy. In unligated structures
247 of HAdV-D26K (PDB 6FJO), Ile324 exhibits a double conformer, occupying the available space (Fig.5C).
248 However, when sialic acid is bound, it is restricted to have a single conformation with the long arm
249 facing away from the sialic acid site, towards the inter-monomer cleft. Ile310 has the opposite
250 orientation and creates an indentation which cradles the methyl group of sialic acid. Tyr312 may
251 further contribute to the hydrophobic cradle. Tyr312 would not normally be considered a hydrophobic
252 residue, but the side chain oxygen faces towards the solvent, where it forms a polar interaction with
253 the C4-OH on sialic acid (Fig.4A), leaving the face of the tyrosine ring exposed to the methyl group
254 which may contribute hydrophobic character to the cradle. This tyrosine behaves in both a polar, and
255 hydrophobic manner at the same time. We suggest that the long arm of Ile324 adopts the sialic acid
256 binding conformation in response to the hydrophobic pressure exerted by sialic acid entering the
257 pocket, minimising the exposed hydrophobic surface when unbound, holding the methyl group
258 between the short arm of Ile310 and the Tyr312 ring, making this an example of induced fit.

259 The double occupancy of Gln348 may indicate a second induced fit mechanism. We observe two
260 possible conformations of Gln348 (Fig.5C). While conformation A does not form any contacts,
261 conformation B forms a polar bond, and water-bridge, with the sialic acid glycerol group. In HAdV-
262 D37K the glycerol group forms only a water-bridge from C7-OH to Ser344, the spatial equivalent of
263 Gln348. Whilst we have not determined the preferred conformation of the sialic acid glycerol group
264 versus Gln348 conformation, we observed that it can form polar contacts with it regardless, while
265 Gln348 is in conformation B (Fig.5C).

266 We suggest Gln348 may be labile until the binding of sialic acid. Upon sialic acid binding Gln348
267 becomes attracted to the charged glycerol group causing it to stabilise in conformation A. This has the
268 effect of "locking" the glycerol side chain in place, which is further restrained by water-bridge contacts
269 to Thr319 and Asn312.

270 Gln348 has greater occupancy in a sialic acid binding conformation (conformation-A) at pH8.0, which
271 corresponds more closely to the physiological conditions in which it would encounter at the cell

surface (Fig.5A). At pH4.0 Gln348 has approximately half occupancy in each conformation (Fig.5B). This implies the possibility of HAdV-D26K having lower sialic acid affinity under more acidic conditions, such as those encountered during endosomal trafficking down the lysosomal pathway.

Therefore, the HAdV-D26K binding pocket to sialic acid is summarised by three synchronous mechanisms. An N-Acetyl anchor comprised of a polar contact to Asn312 stabilised by a water bridge and an induced hydrophobic cradle around the methyl group. An inducible lock, where Gln348 forms a polar contact to the most terminal atoms in the glycerol arm, supported by a network of water bridges. Finally, a network of polar contacts to the carboxyl, C4 oxygen, and nitrogen atoms, which stabilise the pyranose ring.

This interaction in HAdV-D26K is a much more sophisticated binding mechanism compared to HAdV-D19p and the EKC causing viruses. However, the overall pocket topology and several key residues bear similarities. It may be surprising to observe such similarity given the low level of sequence homology HAdV-D26K has to the HAdV-D37K (56.76%, Fig.1A). Other regions, especially the loops, have highly dissimilar sequences. There is a precedent for this within adenovirus, with recombination events being reported in numerous settings(*46*, *47*, *70*, *71*).

It has previously been suggested that many of the species D adenoviruses may have dual sialic acid binding affinity and CAR affinity(*63*). This has been observed in HAdV-D37/64, CAV-2(*19*, *66*), and now HAdV-D26(*62*). Interestingly the species G adenovirus HAdV-G52 has also been observed to bind both CAR and sialic acid, but using two different fiber knob proteins on the same virus and a different mechanism of sialic acid interaction in the knob(*26*), which is shown to bind polysialic acid(*29*). Previous work has proposed CAR may be a receptor for many, if not all, of the species D adenoviruses with variable affinity(*62*, *63*, *72*), and suggest that sialic acid could also be widely utilised(*63*). These findings support that assertion, and adding another species D adenovirus, with low sequence similarity, to the pool of adenoviruses observed to bind both CAR and sialic acid.

Human adenovirus serotypes 43, 27, and 28 fiber-knobs share high sequence homology with HAdV-D26K, sharing the majority of the critical binding residues, and/or having structural homologues at those positions (Supplementary Fig.4). HAdV-D26K is the only species D adenovirus to have a glutamine at position 348 (HAdV-D26K numbering), though many have the shorter, but similarly charged, asparagine at this location, share the serine or similarly charged residue found in HAdV-D37K, or possess an asparagine which could behave similarly to glutamine. However, HAdV-D8 has an uncharged alanine at this position suggesting that a charged residue may not be strictly required for sialic acid binding, though may alter affinity (Supplementary Fig.4).

HAdV-D8K shares an asparagine at the same position as HAdV-D26K which we have shown to form polar and water-bridge contacts to sialic acid (Fig.4A). While this is unique among the classical EKC causing viruses HAdV-D19p/37/64, it is the most common residue at this position in the species D adenoviruses (Supplementary Fig.4).

The HAdV-D26K surface electrostatics are most like those of HAdV-D19pK. HAdV-D19pK is capable of binding sialic acid(*63*), and a limited effect is seen on infection of A549 cell binding after neuraminidase treatment to remove cell surface sialic acid(*69*). HAdV-D19p binding to Chang C (human conjunctival) cells was completely unaffected by neuraminidase treatment, though binding was very low regardless of neuraminidase treatment(*27*). This inability to bind Chang C cells was shown to depend upon a single lysine residue (Lys240) in the apex of the fiber-knob, but distant from the sialic binding pocket, creating a more acidic apical region in the lysine's absence(*73*). HAdV-D26K also lacks a lysine in this position and has the most acidic electrostatic profile observed in this study (Fig.1).

HAdV-D37K, and the identical HAdV-D64K, have been shown to preferentially interact with the sialic acid bearing GD1a glycan on the corneal cell surface, causing EKC(*65*). However, it seems unlikely that a protein capable of trivalent sialic acid binding is completely specific for GD1a, a di-sialylated glycan, given the wide range of available glycan motifs which are di- and tri-sialylated. The GD1a preference may be diminished in HAdV-D19p by the acidic surface caused by the two amino acid substitutions, creating a glycan preference for tissues outside of the eye. Similarly, HAdV-D26K may have a unique glycan preference, driving its tissue tropism towards cells with different glycosylation patterns.

These findings clarify the receptor tropism of HAdV-D26 and build upon the increasingly complex body of knowledge describing species D adenoviruses. The comparison of different sialic acid binding residues suggests greater plasticity regarding the specific residues needed for sialic acid binding than previously thought (Supplementary Fig.4). It seems highly likely that many adenoviruses in species D, and perhaps other species, may interact with sialic acid in this manner. This suggests potential causes of off target infection by species D derived viral vectors. Conversely, investigation of their specific glycan preferences may enable more tissue specific targeting. Knowledge of the sialic acid binding mechanism suggests mutations which may ablate sialic acid interaction, enabling engineering of better restricted tropisms for future virotherapies. This knowledge regarding HAdV-D26 receptor can inform clinical practice in the rare cases of acute HAdV-D26 infection, or in the face of adverse reactions to HAdV-D26 based vaccines, suggesting that sialic acid binding inhibitors, such as Zanamivir, or trivalent sialic acid derivatives(*74*) may make effective anti-HAdV-D26 therapies.

Materials and Methods:

Infectivity assays:

Cells were seeded at a density of 30,000 cells/well in a flat bottomed 96 well cell culture plate and incubated overnight at 37°C to adhere. Cells were washed twice with 200µl of PBS and 50ul of neuraminidase (Sigma-Aldrich, Cat#11080725001) was added to the appropriate wells at a concentration of 50mU/ml, diluted in serum free media, and incubated for 1hr at 37°C. Cells were cooled on ice and washed twice with 200µl of PBS. Green Fluorescent Protein (GFP) expressing, replication incompetent viruses were added to the appropriate wells at a concentration of 2000 or 5000 viral particles per cell, in 100ul of serum free media, at 4°C, and incubated on ice for 1hr. Serum free media alone was added to uninfected control wells. Cells were washed twice with 200µl of cold PBS, complete media added (DMEM, 10% FCS) and incubated for 48hrs at 37°C. Cells were then trypsinised and transferred to a 96 well V-bottom plate, washed twice in 200µl of PBS and fixed in 2% paraformaldehyde for 20mins before wash, and resuspension in 200µl of PBS.

Samples were run in triplicate and analysed by flow cytometry on Attune NxT (ThermoFisher), analysed using FlowJo v10 (FlowJo, LLC), gating sequentially on singlets, cell population, and GFP positive cells. Levels of infection were defined as the percentage of GFP positive cells (%+ve), and/or Total Fluorescence (TF), defined as the percentage of GFP positive cells multiplied by the median fluorescent intensity (MFI) of the GFP positive population. These measures are distinct in that %+ve describes the total proportion of cells infected, and TF describes the total efficiency of transgene delivery.

Amino Acid Sequence Alignments:

Representative whole genomes of HAdV-D64, HAdV-D19p, HAdV-D26, and HAdV-D37 were selected from the National Center for Biotechnology Information (NCBI), the fiber-knob domain amino acid sequences were derived from them, defined as the translated nucleotide sequence of the fiber protein (pIV) from the conserved TLW hinge motif to the protein C-terminus. The fiber-knob domains were aligned using the EMBL-EBI Clustal Omega tool(*75*).

Generation of Recombinant Fiber-Knob protein:

SG13009 *E.coli* harbouring pREP-4 plasmid and pQE-30 expression vector containing the fiber-knob DNA sequence were cultured in 20ml LB broth with 100µg/ml ampicillin and 50µg/ml kanamycin overnight from glycerol stocks made in previous studies(*76–78*). 1L of TB (Terrific Broth, modified, Sigma-Aldrich) containing 100µg/ml ampicillin and 50µg/ml were inoculated with the overnight *E.coli* culture and incubated at 37°C until they reached OD0.6. IPTG was then added to a final concentration

of 0.5mM and the culture incubated at 37°C for 4hrs. Cells were then harvested by centrifugation at 3000g, resuspended in lysis buffer (50mM Tris, pH8.0, 300mM NaCl, 1% (v/v) NP40, 1mg/ml Lysozyme, 1mM β-mercaptoethanol), and incubated at room temperature for 30mins. Lysate was clarified by centrifugation at 30,000g for 30mins and filtered through a 0.22µm syringe filter (Millipore, Abingdon, UK). Clarified lysate was then loaded onto a 5ml HisTrap FF nickel affinity chromatography column (GE) at 2.0ml/min and washed with 5 column volumes into elution buffer A (50mM Tris [pH8.0], 300mM NaCl, 1mM β-mercaptoethanol). Protein was eluted by 30min gradient elution from buffer A to B (buffer A + 400mM Imidazole). Fractions were analysed by reducing SDS-PAGE, and Fiber-knob containing fractions further purified using a superdex 200 10/300 size exclusion chromatography column (GE) in crystallisation buffer (10 mM Tris [pH 8.0] and 30 mM NaCl). Fractions were analysed by SDS-PAGE and pure fractions concentrated by centrifugation in Vivaspin 10,000 MWCO (Sartorius, Goettingen, Germany) proceeding crystallisation.

Crystallisation and structure determination.

Protein samples were purified into crystallisation buffer (10 mM TRIS [pH 8.0] and 30 mM NaCl). The final protein concentration was approximately 10 mg/ml. Commercial crystallisation screen solutions were dispensed into 96-well plates using an Art-Robbins Instruments Griffon dispensing robot (Alpha Biotech, Ltd), in sitting-drop vapour-diffusion format. Drops containing 200nl of screen solution and 200nl of protein solution were equilibrated against a reservoir of 60µl crystallisation solution. The plates were sealed and incubated at 18°C.

Crystals of HAdV-D26K appeared in PACT Premier condition B01 and B04 (0.1 M MIB [Malonic acid, Imidazole, Boric acid], 25 % w/v PEG 1500, pH4.0 and pH8.0 respectively), within 1 to 7 days. Crystals were then soaked in reservoir solution containing N-Acetylneuraminic acid (Neu5Ac, Sigma-Aldrich Cat#A2388) at a final concentration of 10mM and incubated overnight prior to harvest. Crystals were cryoprotected with reservoir solution to which ethylene glycol was added at a final concentration of 25%. Crystals were harvested in thin plastic loops and stored in liquid nitrogen for transfer to the synchrotron. Data were collected at Diamond Light Source beamline I04, running at a wavelength of 0.9795Å. During data collection, crystals were maintained in a cold air stream at 100°K. Dectris Pilatus 6M detectors recorded the diffraction patterns, which were analysed and reduced with XDS(*79*), Xia2, DIALS(*80*), and Autoproc(*81*). Scaling and merging data was completed with Pointless, Aimless and Truncate from the CCP4 package(*82*). Structures were solved with PHASER, COOT was used to correct the sequences and adjust the models, REFMAC5 was used to refine the structures and calculate maps. Graphical representations were prepared with PyMOL(*83*). Reflection data and final models were

399 deposited in the PDB database with accession codes: 6QU6, 6QU8, and 6FJO. Full crystallographic
400 refinement statistics are given in Supplementary Table 1

401 Calculation of electrostatic surface potentials and isoelectric points:

402 HAdV-D37, HAdV-D19p, and HAdV-D26 used PDB IUXA, PDB 1UXB, and PDB 6QU8, respectively, as the
403 input. HAdV-D8 was calculated using a homology model, generated as described below, for input.

404 The PDB2PQR server (V 2.1.1)(*84*) was used assign charge and radius parameters using the PARSE
405 forcefield, and assigned protonation states using PROPKA, at pH7.35. APBS(*85*) was used to calculate
406 electrostatic surface potentials, and the map output was visualised in PyMol(*83*).

407 Homology modelling of Adenovirus serotype 8

408 The I-TASSER protein structure and function prediction server(*86–88*) was used to generate a
409 homology model of HAdV-D8 based on the published sequence of HAdV-D8(*42*), using the published
410 structure of it's closest relative (by sequence identity), HAdV-D19p(*63*). The resultant monomer was
411 then copied three times, using the HAdV-19p trimer as a template, and the monomers aligned in
412 PyMol(*83*) so as to generate a model of the complete HAdV-D8K trimer.

413 Acknowledgements

414 A.T.B is supported by a Tenovus Cancer Care PhD studentship to A.L.P (project reference
415 PhD2015/L13). J. A. D is supported by a Cancer Research UK Biotherapeutics Drug Discovery Project
416 Award to A.L.P. (project reference C52915/A23946). A.L.P and P.R are funded by Higher Education
417 Funding Council for Wales.

418 The authors wish to acknowledge the Diamond Light Source for beamtime (proposal mx14843), and
419 the staff of beamline I04 for assistance with diffraction data collection.

1. T. Lion, Adenovirus Infections in Immunocompetent and Immunocompromised Patients. *Clin. Microbiol. Rev.* **27**, 441–462 (2014).

2. Adenovirus Clinical Presentation: History, Physical, Complications, (available at https://emedicine.medscape.com/article/211738-clinical).

3. D. Garcia-Zalisnak, C. Rapuano, J. D. Sheppard, A. R. Davis, Adenovirus Ocular Infections: Prevalence, Pathology, Pitfalls, and Practical Pointers. *Eye Contact Lens*. **44 Suppl 1**, S1–S7 (2018).

4. R. Civljak *et al.*, Viral pathogens associated with acute respiratory illness in hospitalized adults and elderly from Zagreb, Croatia, 2016 to 2018. *J. Med. Virol.* (2019), doi:10.1002/jmv.25437.

5. C. A. Pfortmueller *et al.*, Severe acute respiratory distress syndrome (ARDS) induced by human adenovirus B21: Report on 2 cases and literature review. *J. Crit. Care*. **51**, 99–104 (2019).

6. K. Kosulin *et al.*, *Clin. Microbiol. Infect.*, in press, doi:10.1016/j.cmi.2015.12.013.

7. C. M. Robinson *et al.*, Computational analysis and identification of an emergent human adenovirus pathogen implicated in a respiratory fatality. *Virology*. **409**, 141–147 (2011).

8. M. E. Killerby *et al.*, Respiratory Illness Associated With Emergent Human Adenovirus Genome Type 7d, New Jersey, 2016-2017. *Open Forum Infect. Dis.* **6**, ofz017 (2019).

9. R. N. Potter, J. A. Cantrell, C. T. Mallak, J. C. Gaydos, Adenovirus-associated Deaths in US Military during Postvaccination Period, 1999–2010. *Emerg. Infect. Dis.* **18**, 507–509 (2012).

10. S. A. Kolavic-Gray *et al.*, Large Epidemic of Adenovirus Type 4 Infection among Military Trainees: Epidemiological, Clinical, and Laboratory Studies. *Clin. Infect. Dis.* **35**, 808–818 (2002).

11. S. Khanal, P. Ghimire, A. S. Dhamoon, The Repertoire of Adenovirus in Human Disease: The Innocuous to the Deadly. *Biomedicines*. **6**, 30 (2018).

12. S. H. Norris, T. C. Butler, N. Glass, R. Tran, Fatal Hepatic Necrosis Caused by Disseminated Type 5 Adenovirus Infection in a Renal Transplant Recipient. *Am. J. Nephrol.* **9**, 101–105 (1989).

13. Y. M. A. Al Qurashi, M. A. Alkhalaf, L. Lim, M. Guiver, R. J. Cooper, Sequencing and phylogenetic analysis of the hexon, fiber, and penton regions of adenoviruses isolated from AIDS patients. *J. Med. Virol.* **84**, 1157–1165 (2012).

14. J. C. Hierholzer, Adenoviruses in the immunocompromised host. *Clin. Microbiol. Rev.* **5**, 262–274 (1992).

15. A. M. Q. King, M. J. Adams, E. B. Carstens, E. J. Lefkowitz, Eds., in *Virus Taxonomy* (Elsevier, San Diego, 2012; http://www.sciencedirect.com/science/article/pii/B9780123846846000094), pp. 125–141.

16. International Committee on Taxonomy of Viruses, *Virus taxonomy: classification and nomenclature of viruses : sixth report of the International Committee on Taxonomy of Viruses* (Springer-Verlag, Wien ; New York, 1995), *Archives of virology. Supplementum ; 10*.

17. HAdV Working Group, http://hadvwg.gmu.edu/, (available at http://hadvwg.gmu.edu/).

18. J. M. Bergelson *et al.*, Isolation of a Common Receptor for Coxsackie B Viruses and Adenoviruses 2 and 5. *Science*. **275**, 1320–1323 (1997).

19. E. Seiradake, H. Lortat-Jacob, O. Billet, E. J. Kremer, S. Cusack, 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**, 33704–33716 (2006).

20. A. Gaggar, D. M. Shayakhmetov, A. Lieber, CD46 is a cellular receptor for group B adenoviruses. *Nat. Med.* **9**, 1408–1412 (2003).

21. C. Fleischli *et al.*, Species B adenovirus serotypes 3, 7, 11 and 35 share similar binding sites on the membrane cofactor protein CD46 receptor. *J. Gen. Virol.* **88**, 2925–2934 (2007).

22. M. Marttila *et al.*, CD46 is a cellular receptor for all species B adenoviruses except types 3 and 7. *J. Virol.* **79**, 14429–14436 (2005).

23. H. Wang *et al.*, Desmoglein 2 is a receptor for adenovirus serotypes 3, 7, 11 and 14. *Nat. Med.* **17**, 96–104 (2011).

24. E. Vassal-Stermann *et al.*, Mapping of Adenovirus of serotype 3 fibre interaction to desmoglein 2 revealed a novel 'non-classical' mechanism of viral receptor engagement. *Sci. Rep.* **8** (2018), doi:10.1038/s41598-018-26871-x.

25. E. Vassal-Stermann *et al.*, Cryo-EM structure of adenovirus type 3 fibre with desmoglein 2 shows a novel mode of receptor engagement. *bioRxiv*, 471383 (2018).

26. A. Lenman *et al.*, Human Adenovirus 52 Uses Sialic Acid-containing Glycoproteins and the Coxsackie and Adenovirus Receptor for Binding to Target Cells. *PLOS Pathog.* **11**, e1004657 (2015).

27. N. Arnberg, P. Pring-Åkerblom, G. Wadell, Adenovirus Type 37 Uses Sialic Acid as a Cellular Receptor on Chang C Cells. *J. Virol.* **76**, 8834–8841 (2002).

28. N. Arnberg, K. Edlund, A. H. Kidd, G. Wadell, Adenovirus Type 37 Uses Sialic Acid as a Cellular Receptor. *J. Virol.* **74**, 42–48 (2000).

29. A. Lenman *et al.*, Polysialic acid is a cellular receptor for human adenovirus 52. *Proc. Natl. Acad. Sci.*, 201716900 (2018).

30. P. Abbink *et al.*, Comparative Seroprevalence and Immunogenicity of Six Rare Serotype Recombinant Adenovirus Vaccine Vectors from Subgroups B and D. *J. Virol.* **81**, 4654–4663 (2007).

31. M. Aste-Amézaga *et al.*, Quantitative adenovirus neutralization assays based on the secreted alkaline phosphatase reporter gene: application in epidemiologic studies and in the design of adenovector vaccines. *Hum. Gene Ther.* **15**, 293–304 (2004).

32. S. J. Moyo *et al.*, Prevalence and molecular characterisation of human adenovirus in diarrhoeic children in Tanzania; a case control study. *BMC Infect. Dis.* **14**, 666 (2014).

33. A. R. Thorner *et al.*, Age Dependence of Adenovirus-Specific Neutralizing Antibody Titers in Individuals from Sub-Saharan Africa. *J. Clin. Microbiol.* **44**, 3781–3783 (2006).

493 34. M. Pauly *et al.*, High prevalence and diversity of species D adenoviruses (HAdV-D) in human
494 populations of four Sub-Saharan countries. *Virol. J.* **11**, 25 (2014).

495 35. R. S. Dakin, A. L. Parker, C. Delles, S. A. Nicklin, A. H. Baker, Efficient transduction of primary
496 vascular cells by the rare adenovirus serotype 49 vector. *Hum. Gene Ther.* **26**, 312–319 (2015).

497 36. P. A. Phillips, G. B. Harnett, M. M. Gollow, Adenovirus type 19 and a closely related new serotype
498 in genital infection. *Br. J. Vener. Dis.* **58**, 131–132 (1982).

499 37. P. D. Swenson, M. S. Lowens, C. L. Celum, J. C. Hierholzer, Adenovirus types 2, 8, and 37
500 associated with genital infections in patients attending a sexually transmitted disease clinic. *J.*
501 *Clin. Microbiol.* **33**, 2728–2731 (1995).

502 38. K. Aoki, Y. Tagawa, A Twenty-One Year Surveillance of Adenoviral Conjunctivitis in Sapporo,
503 Japan. *Int. Ophthalmol. Clin.* **42**, 49–54 (2002).

504 39. P. R. Kinchington, E. G. Romanowski, Y. Jerold Gordon, Prospects for adenovirus antivirals. *J.*
505 *Antimicrob. Chemother.* **55**, 424–429 (2005).

506 40. Z. Lei *et al.*, Outbreaks of epidemic keratoconjunctivitis caused by human adenovirus type 8 in
507 the Tibet Autonomous Region of China in 2016. *PLoS ONE*. **12** (2017),
508 doi:10.1371/journal.pone.0185048.

509 41. E. Hage *et al.*, Molecular phylogeny of a novel human adenovirus type 8 strain causing a
510 prolonged, multi-state keratoconjunctivitis epidemic in Germany. *Sci. Rep.* **7** (2017),
511 doi:10.1038/srep40680.

512 42. H. Kaneko *et al.*, Analysis of the complete genome sequence of epidemic keratoconjunctivitis-
513 related human adenovirus type 8, 19, 37 and a novel serotype. *J. Gen. Virol.* **90**, 1471–1476
514 (2009).

515 43. J. C. de Jong *et al.*, Adenovirus 37: identification and characterization of a medically important
516 new adenovirus type of subgroup D. *J. Med. Virol.* **7**, 105–118 (1981).

517 44. T. Ariga *et al.*, Five New Genome Types of Adenovirus Type 37 Caused Epidemic
518 Keratoconjunctivitis in Sapporo, Japan, for More Than 10 Years. *J. Clin. Microbiol.* **43**, 726–732
519 (2005).

520 45. C. M. Robinson *et al.*, Human adenovirus type 19: genomic and bioinformatics analysis of a
521 keratoconjunctivitis isolate. *Virus Res.* **139**, 122–126 (2009).

522 46. X. Zhou *et al.*, Analysis of human adenovirus type 19 associated with epidemic
523 keratoconjunctivitis and its reclassification as adenovirus type 64. *Invest. Ophthalmol. Vis. Sci.*
524 **53**, 2804–2811 (2012).

525 47. M. P. Walsh *et al.*, Evidence of Molecular Evolution Driven by Recombination Events Influencing
526 Tropism in a Novel Human Adenovirus that Causes Epidemic Keratoconjunctivitis. *PLOS ONE*. **4**,
527 e5635 (2009).

528 48. K. Aoki, N. Kitaichi, R. Hinokuma, S. Ohno, Characteristics of Epidemic Keratoconjunctivitis due
529 to Novel Adenovirus Types. **1**, 3 (2017).

49. K. Matsuura *et al.*, Human adenoviral type 54 keratoconjunctivitis accompanied by stellate keratitis and keratic precipitates: two cases. *BMC Ophthalmol.* **19**, 7 (2019).

50. G. Huang *et al.*, Outbreak of Epidemic Keratoconjunctivitis Caused by Human Adenovirus Type 56, China, 2012. *PLOS ONE*. **9**, e110781 (2014).

51. R. J. Stanton, B. P. McSharry, M. Armstrong, P. Tomasec, G. W. G. Wilkinson, Re-engineering adenovirus vector systems to enable high-throughput analyses of gene function. *BioTechniques*. **45**, 659–662, 664–668 (2008).

52. H. Uusi-Kerttula, S. Hulin-Curtis, J. Davies, A. L. Parker, Oncolytic Adenovirus: Strategies and Insights for Vector Design and Immuno-Oncolytic Applications. *Viruses*. **7**, 6009–6042 (2015).

53. D. Majhen *et al.*, Adenovirus-Based Vaccines for Fighting Infectious Diseases and Cancer: Progress in the Field. *Hum. Gene Ther.* **25**, 301–317 (2014).

54. H. Chen *et al.*, Adenovirus-Based Vaccines: Comparison of Vectors from Three Species of Adenoviridae. *J. Virol.* **84**, 10522–10532 (2010).

55. D. H. Barouch *et al.*, Evaluation of a mosaic HIV-1 vaccine in a multicentre, randomised, double-blind, placebo-controlled, phase 1/2a clinical trial (APPROACH) and in rhesus monkeys (NHP 13-19). *The Lancet*. **392**, 232–243 (2018).

56. L. R. Baden *et al.*, First-in-Human Evaluation of a Hexon Chimeric Adenovirus Vector Expressing HIV-1 Env (IPCAVD 002). *J. Infect. Dis.* **210**, 1052–1061 (2014).

57. F. Cox *et al.*, Adenoviral vector type 26 encoding Zika virus (ZIKV) M-Env antigen induces humoral and cellular immune responses and protects mice and nonhuman primates against ZIKV challenge. *PLOS ONE*. **13**, e0202820 (2018).

58. Long-term Safety Follow-up of Participants Exposed to the Candidate Ebola Vaccines Ad26.ZEBOV and/or MVA-BN-Filo - Full Text View - ClinicalTrials.gov, (available at https://clinicaltrials.gov/ct2/show/NCT02661464?term=adenovirus&recrs=a&cond=ebola&rank=1).

59. T. W. Geisbert *et al.*, Recombinant Adenovirus Serotype 26 (Ad26) and Ad35 Vaccine Vectors Bypass Immunity to Ad5 and Protect Nonhuman Primates against Ebolavirus Challenge. *J. Virol.* **85**, 4222–4233 (2011).

60. Trial of the Safety and Immunogenicity of an Oral, Replicating Ad26 Vectored HIV-1 Vaccine - Full Text View - ClinicalTrials.gov, (available at https://clinicaltrials.gov/ct2/show/NCT02366013).

61. A Shedding Study of Adenovirus Serotype 26 Based Respiratory Syncytial Virus Pre-fusion F Protein (Ad26.RSV.preF) Vaccine in Adults - Full Text View - ClinicalTrials.gov, (available at https://clinicaltrials.gov/ct2/show/NCT03795441).

62. A. T. Baker *et al.*, Diversity within the adenovirus fiber knob hypervariable loops influences primary receptor interactions. *bioRxiv*, 406819 (2018).

63. W. P. Burmeister, D. Guilligay, S. Cusack, G. Wadell, N. Arnberg, Crystal Structure of Species D Adenovirus Fiber Knobs and Their Sialic Acid Binding Sites. *J. Virol.* **78**, 7727–7736 (2004).

64. S. M. Arcasoy *et al.*, MUC1 and Other Sialoglycoconjugates Inhibit Adenovirus-mediated Gene Transfer to Epithelial Cells. *Am. J. Respir. Cell Mol. Biol.* **17**, 422–435 (1997).

65. E. C. Nilsson *et al.*, The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. *Nat. Med.* **17**, 105–109 (2011).

66. E. Seiradake *et al.*, The cell adhesion molecule "CAR" and sialic acid on human erythrocytes influence adenovirus in vivo biodistribution. *PLoS Pathog.* **5**, e1000277 (2009).

67. A. K. Singh *et al.*, Structure and Sialyllactose Binding of the Carboxy-Terminal Head Domain of the Fibre from a Siadenovirus, Turkey Adenovirus 3. *PloS One*. **10**, e0139339 (2015).

68. M.-C. Amoureux *et al.*, Polysialic Acid Neural Cell Adhesion Molecule (PSA-NCAM) is an adverse prognosis factor in glioblastoma, and regulates olig2 expression in glioma cell lines. *BMC Cancer*. **10**, 91 (2010).

69. N. Arnberg, A. H. Kidd, K. Edlund, F. Olfat, G. Wadell, Initial Interactions of Subgenus D Adenoviruses with A549 Cellular Receptors: Sialic Acid versus αvIntegrins. *J. Virol.* **74**, 7691–7693 (2000).

70. C. M. Robinson *et al.*, Molecular evolution of human adenoviruses. *Sci. Rep.* **3**, 1812 (2013).

71. A. N. Lukashev, O. E. Ivanova, T. P. Eremeeva, R. D. Iggo, Evidence of frequent recombination among human adenoviruses. *J. Gen. Virol.* **89**, 380–388 (2008).

72. I. Kirby *et al.*, 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 (2001).

73. S. Huang, V. Reddy, N. Dasgupta, G. R. Nemerow, A Single Amino Acid in the Adenovirus Type 37 Fiber Confers Binding to Human Conjunctival Cells. *J. Virol.* **73**, 2798–2802 (1999).

74. S. Spjut *et al.*, A Potent Trivalent Sialic Acid Inhibitor of Adenovirus Type 37 Infection of Human Corneal Cells. *Angew. Chem. Int. Ed.* **50**, 6519–6521 (2011).

75. F. Sievers *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539–539 (2011).

76. L. Coughlan *et al.*, Retargeting Adenovirus Serotype 48 Fiber Knob Domain by Peptide Incorporation. *Hum. Gene Ther.* **25**, 385–394 (2014).

77. H. Uusi-Kerttula *et al.*, Pseudotyped αvβ6 integrin-targeted adenovirus vectors for ovarian cancer therapies. *Oncotarget*. **7**, 27926–27937 (2016).

78. L. Coughlan *et al.*, In Vivo Retargeting of Adenovirus Type 5 to αvβ6 Integrin Results in Reduced Hepatotoxicity and Improved Tumor Uptake following Systemic Delivery. *J. Virol.* **83**, 6416–6428 (2009).

79. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

80. G. Winter *et al.*, DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 85–97 (2018).

603 81. C. Vonrhein *et al.*, Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 293–302 (2011).

605 82. E. J. Dodson, M. Winn, A. Ralph, in *Methods in Enzymology* (Academic Press, 1997; http://www.sciencedirect.com/science/article/pii/S0076687997770344), vol. 277 of *Macromolecular Crystallography Part B*, pp. 620–633.

608 83. Schrödinger, LLC., The PyMOL Molecular Graphics System, Version 2.0.

609 84. T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665-667 (2004).

612 85. N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10037–10041 (2001).

615 86. A. Roy, A. Kucukural, Y. Zhang, I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).

617 87. J. Yang *et al.*, The I-TASSER Suite: protein structure and function prediction. *Nat. Methods*. **12**, 7–8 (2015).

619 88. Y. Zhang, I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. **9**, 40 (2008).

620

Figure 1: HAdV-D26K forms a local basic area in the apical depression to facilitate sialic acid binding despite net positive surface potential. HAdV-D26K has low (56.76%) sequence identity with fiber-knobs known to bind sialic acid by a similar mechanism, and an acidic isoelectric point (**A**). The electrostatic potential surfaces of HAdV-D8K (**B**), HAdV-D64/37 (**C**), and HAdV-D19p (**D**) fiber-knobs are highly basic, especially about the central depression about the 3-fold axis. HAdV-D26 fiber-knob is less basic overall but maintains positive potential in the central depression (**E**). Surfaces are displayed at ±10mV.

Figure 2: HAdV-D26K shares key binding residues with sialic acid utilising adenoviruses and exploits sialic acid to infect cells. Sequence alignment of HAdV-D26K shows conservation of key binding residues with known sialic acid binding adenoviruses (**A**). Residues highlighted in Red form polar contacts with sialic acid, green contact sialic acid via water-bridge, a black underline indicates both direct and water-bridge contacts, blue indicates hydrophobic contacts. Neuraminidase treatment does not reduce the ability of HAdV-D5/B35K (**B**) or HAdV-C5 (**C**) to infect SKOV-3 (ovarian adenocarcinoma), BT-20 (breast carcinoma), or MDA-231 (metastatic breast adenocarcinoma) cells, while HAdV-D5/D26K (**D**) is significantly inhibited. n=3 biological replicates, error = ±SD.

Figure 3: Sialic acid binds in the apical depression of adenovirus 26 fiber-knob protein. Sialic acid (orange) is seen to bind in 3 locations in the apical depression of the HAdV-D26 fiber-knob, bridging between monomers (shades of blue) of the trimeric assembly (**A**). The map shows clear density for a ligand (**B**), which is best described by a double conformer of sialic acid (**C**). Crystallisation statistics in supplementary table 1, 2FoFc map (blue mesh, σ=1.5), FoFc (Green mesh, σ=3.0).

Figure 4: HAdV-D26K forms a complex interaction network of hydrophobic and electrostatic interactions with sialic acid. Sialic acid (orange) is seen to bind HAdV-D26 (**A**) and HAdV-D37 (**B**) through a network of polar contacts (red dashes) and hydrogen bonds (blue dashes). The interaction is stabilised by hydrophobic interactions (red regions on white surface) with the N-Acetyl CH_3 group, but different residues in HAdV-D26 (**C**) and HAdV-D37 (**D**). Waters are shown as cyan spheres, residues forming comparable contacts in HAdV-D26 and HAdV-D37 are shown as blue sticks, other residues are shown as green sticks. Oxygen and nitrogen are seen in red and blue, respectively.

Figure 5: HAdV-D26K affects an induced fit mechanism in sialic acid binding. HAdV-D26K residue Gln348 can occupy multiple conformations, with a greater preference for conformation A (capable of forming a polar contact with the glycerol arm of sialic acid) at pH8.0 (**A**) than at pH4.0 (**B**). Ile324 has two conformations when HAdV-D26K is unliganded (**C**, PDB 6FJO). However, upon sialic acid binding the Ile324 adopts a single confirmation creating a hydrophobic indentation around the N-Acetyl methyl group bounded by Ile324,Ile310, and the ring of Tyr312 (**D**).

Supplementary Figure 1: Sialic acid forms a stable interaction with HAdV-D26K at both pH4.0 (PDB 6QU6) and pH8.0 (PDB 6QU8). The omit map at pH4.0 (**A**) shows density for a small molecule ligand, which can be best modelled by a sialic acid double conformer (**B**). The same is true at pH8.0 (**C**), but the preferred conformations of the glycerol group are different (**D**). Crystallisation statistics in supplementary table 1, 2FoFc map (blue mesh, σ=1.5), FoFc (Green mesh, σ=3.0).

Supplementary Figure 2: Structure of sialic acid (N-Acetylneuraminic acid, Neu5Ac) in a biologically relevant conformation. Viewing the Neu5Ac face on with glycerol and N-Acetyl groups labelled in red boxes, and the carbons numbered (**A**). Side on the carboxyl group (red box) is seen in the axial conformation with the C2 OH group planar to the ring (**B**).

Supplementary Figure 3: HAdV-D26K forms a similar interaction with sialic acid at both pH4.0 (PDB 6QU6) and pH8.0 (PDB 6QU8) through a combination of polar, water-bridge, and hydrophobic interactions. At pH4.0 sialic acid forms numerous polar contacts to charged side chains in HAdV-D26K (**A**), similar contacts are seen at pH8.0 and exhibits a lysine double conformer (**B**). At pH4.0 sialic acid forms several water bridges stabilising the interaction of the glycerol group (**C**), the same bridges are seen at pH8.0 with the addition of a water-bridge contact on the carboxyl group not seen at pH4.0 (**D**). A hydrophobic interface is formed around the N-Acetyl methyl group which appears to be similar and stable at pH4.0 (**E**) and pH8.0 (**F**). Polar bonds to residues are shown as red dashes, water-bridge contacts as blue dashes, and waters as cyan spheres. Sialic acid is shown in orange, polar HAdV-D26K residues as green sticks, and purely hydrophobic residues as red sticks. The HAdV-D26 surface is shown in white with hydrophobic regions in red. Oxygen and nitrogen atoms are coloured red and blue, respectively.

Supplementary Figure 4: Species D adenoviruses conserve known sialic acid binding residues. Sequence alignment of species D adenovirus fiber-knob proteins. Known HAdV-D26 and HAdV-D37/64/19p residues forming contacts with sialic acid are highlighted in black and red respectively. Homologous residues are coloured similarly to the virus which they share the residue with. HAdV-D8 residues at known sialic acid binding locations which are dissimilar to HAdV-D26/37 are highlighted in blue, as are homologous residues in other viruses. Names utilise the short nomenclature, all are human species D adenoviruses. Numbering is for HAdV-D26K.

Figure 1

A

Fiber-Knob Serotype	Sequence Used (Accession)	Sequence Identity to HAdV-D37	Theoretical $_{p}I$
HAdV-D8	AB448767.1	85.16%	9.03
HAdV-D64	AB448772.1	100.00%	9.14
HAdV-D19p	JQ326209.1	98.90%	8.64
HAdV-D26	EF153474.1	56.76%	6.49
HAdV-D37	AB448776.1	100.00%	9.14

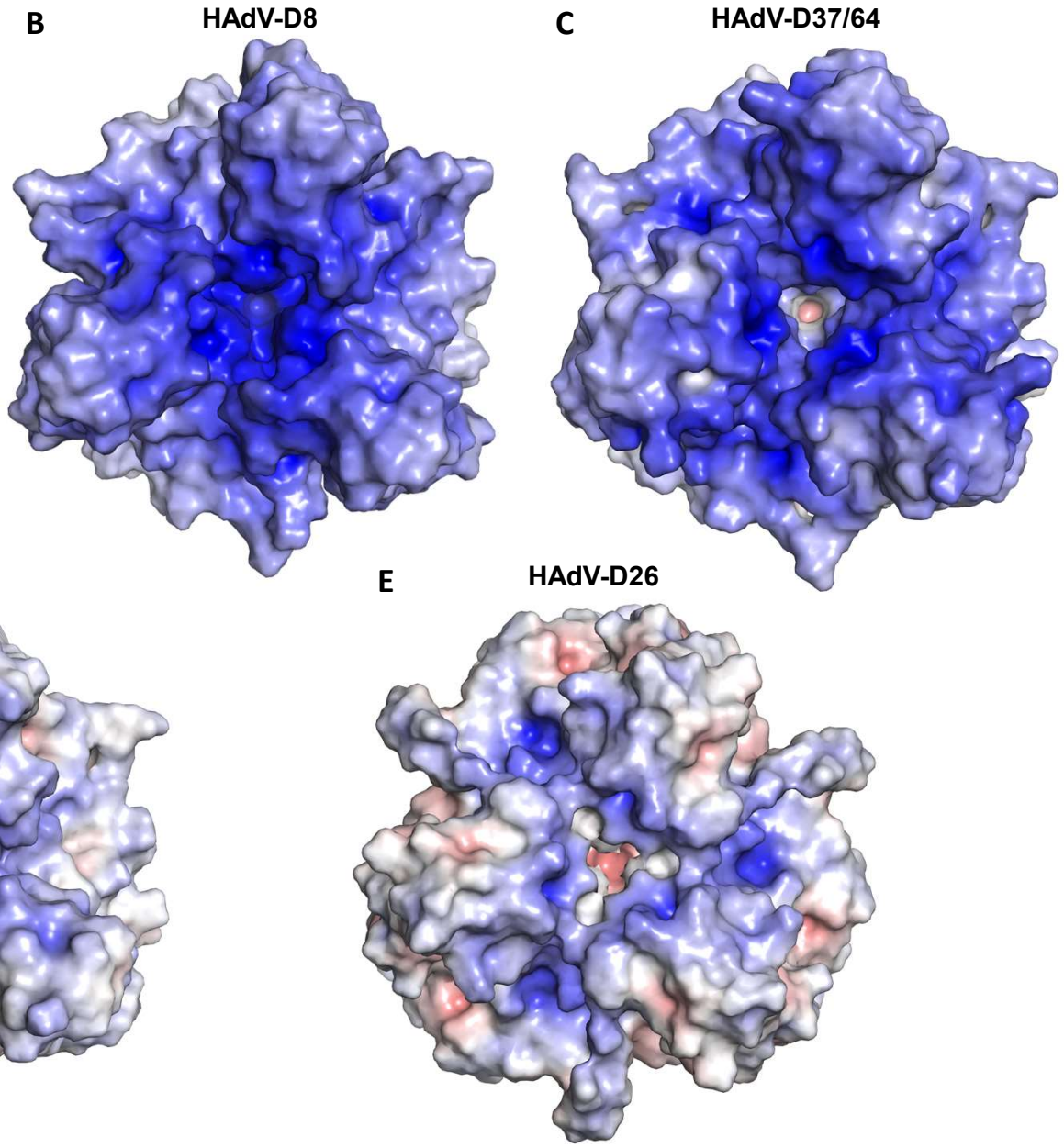


Figure 2

A

```
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
HAdV-D26   188  TLWTTPDTSPNCKMSTEKDSKLTLTLTKCGSQVLGNVSLLAVTGEYHQMTATTK---KDVKISLLFDENG
HAdV-D8    181  TLWTTPDTSPNCRIDQDKDSKLTLVLTKCGSQILANVSLIVVAGRYKIINNNTNPALKGFTIKLLFDKNG
HAdV-D37   185  TLWTTPDTSPNCTIAQDKDSKLTLVLTKCGSQILANVSLIVVAGKYHIINNKTNPKIKSFTIKLLFNKNG
HAdV-D19p  185  TLWTTPDTSPNCTIAQDKDSKLTLVLTKCGSQILANVSLIVVAGKYHIINNKTNPEIKSFTIKLLFNKNG

                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
HAdV-D26   255  ILLPSSSLSKDYWNYRSDDSIVSQKYNNAVPFMPNLTAYPKPSAQNAKNYSRTKIISNVYLGALTYQPVI
HAdV-D8    251  VLMESSNLGKSYWNFRNQNSIMSTAYEKAIGFMPNLVAYPKPTTG-SKKYARDIVYGNIYLGGKPHQPVT
HAdV-D37   255  VLLDNSNLGKAYWNFRSGNSNVSTAYEKAIGFMPNLVAYPKPSN--SKKYARDIVYGTIYLGGKPDQPAV
HAdV-D19p  255  VLLDNSNLGKAYWNFRSGNSNVSTAYEKAIGFMPNLVAYPKPSN--SKKYARDIVYGTIYLGGKPDQPAV

                    ....|....|....|....|....|....|....|....|....|....|....|
HAdV-D26   328  ITIAFNQETENGCAYSITFTFTWQKDYSAQQFDVTSFTFSYLTQENKDKD
HAdV-D8    320  IKTTFNQE--TGCEYSITFDFSWAKTYVNVEFETTSFTFSYIAQE-----
HAdV-D37   323  IKTTFNQE--TGCEYSITFNFSWSKTYENVEFETTSFTFSYIAQE-----
HAdV-D19p  323  IKTTFNQE--TGCEYSITFDFSWSKTYENVEFETTSFTFSYIAQE-----
```

Note: HAdV-D37K is identical to HAdV-D64K

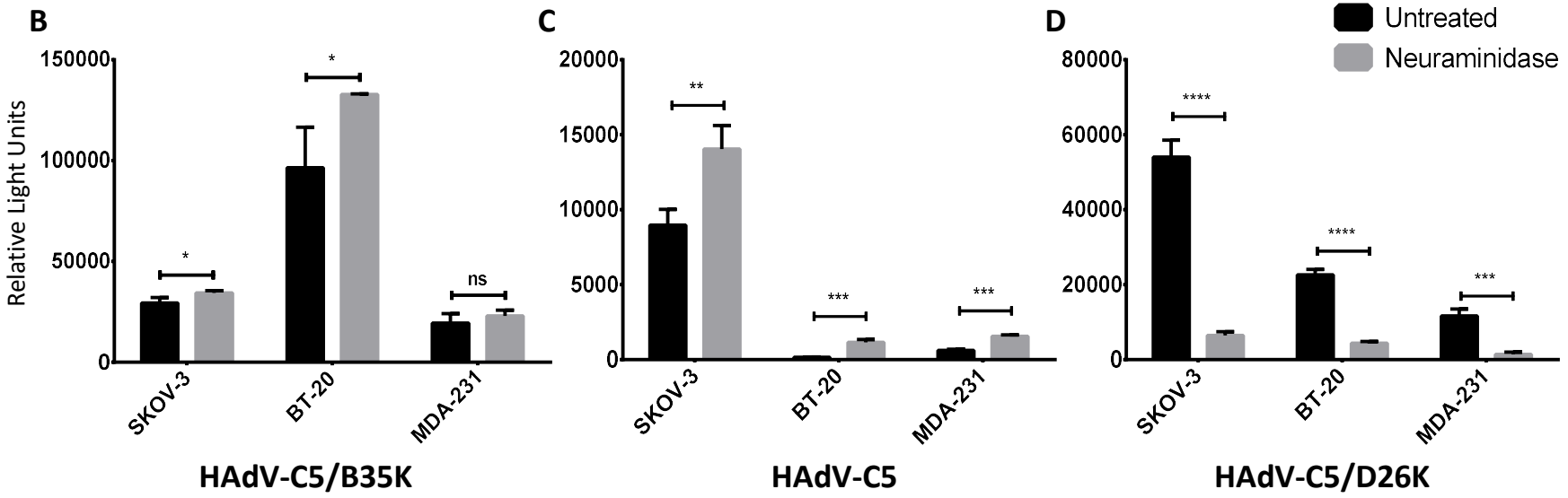


Figure 3

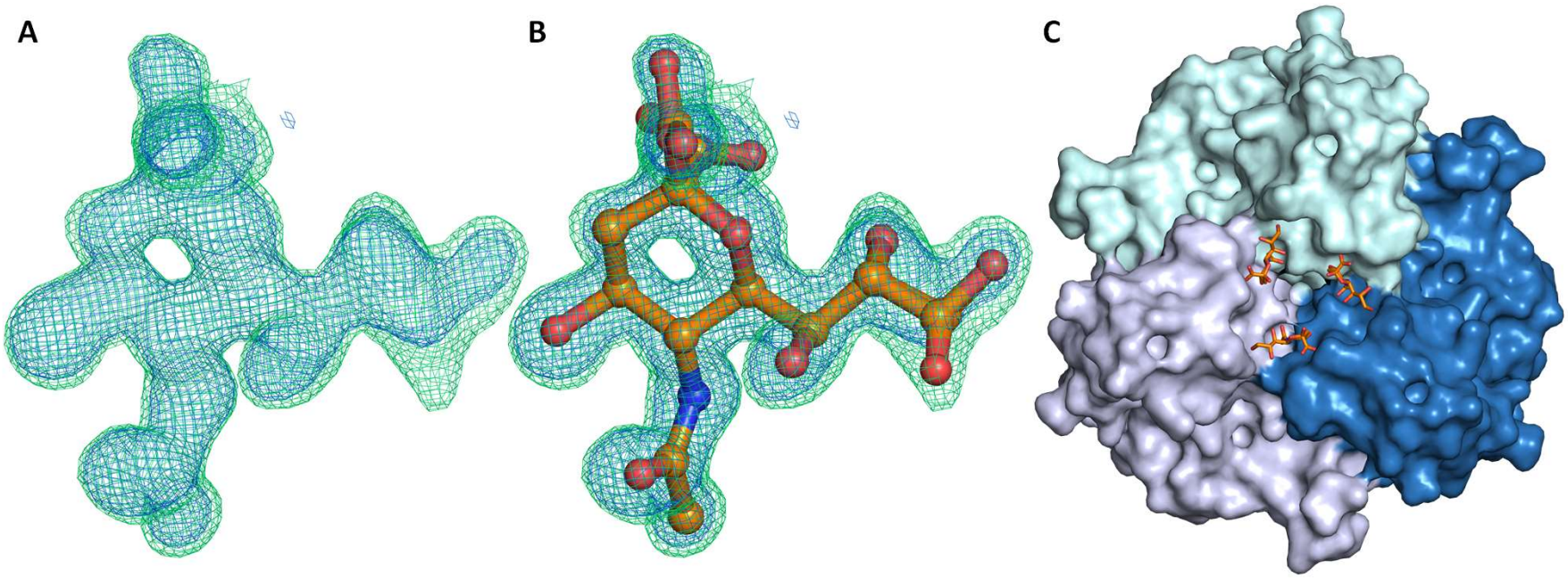


Figure 4

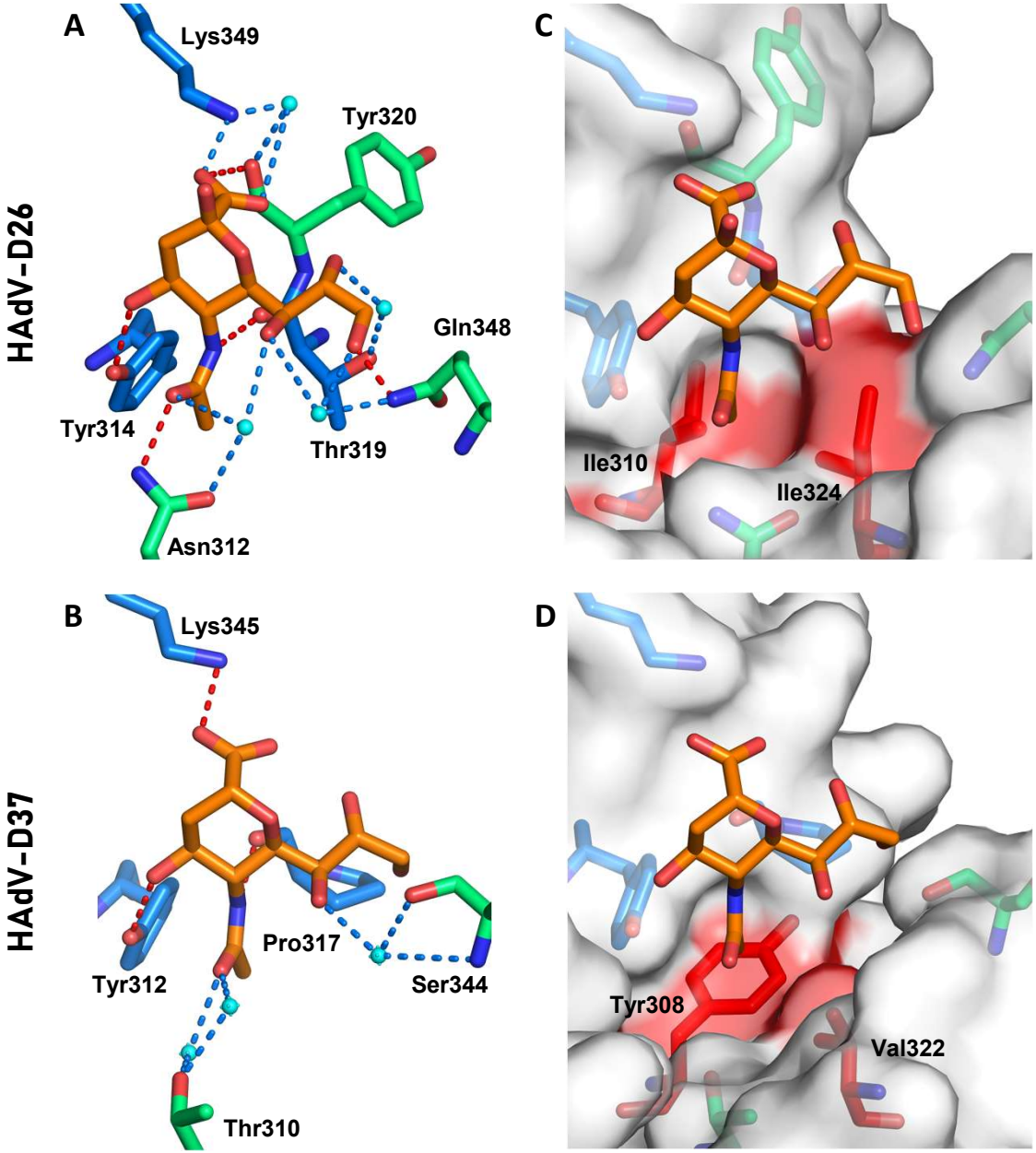
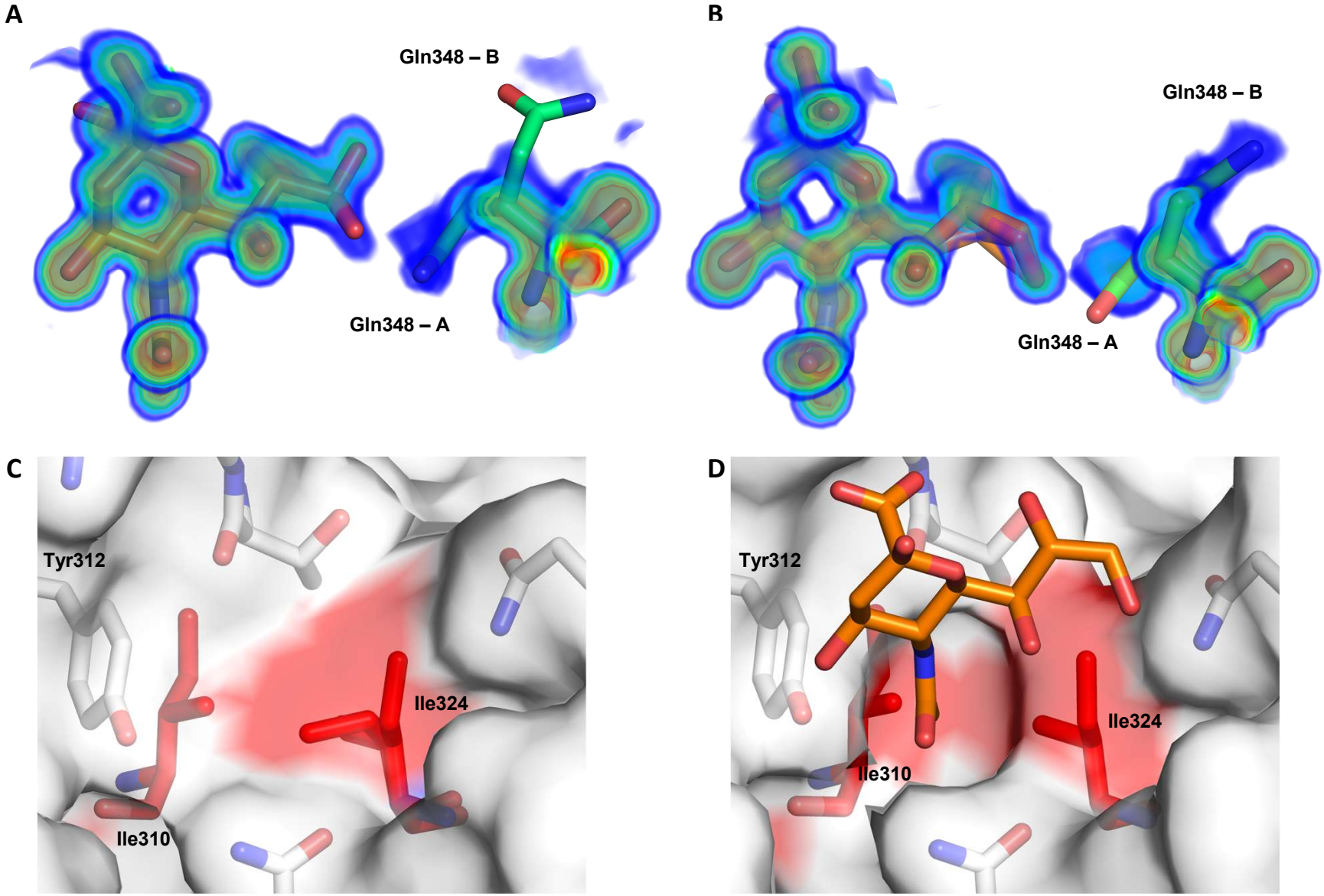
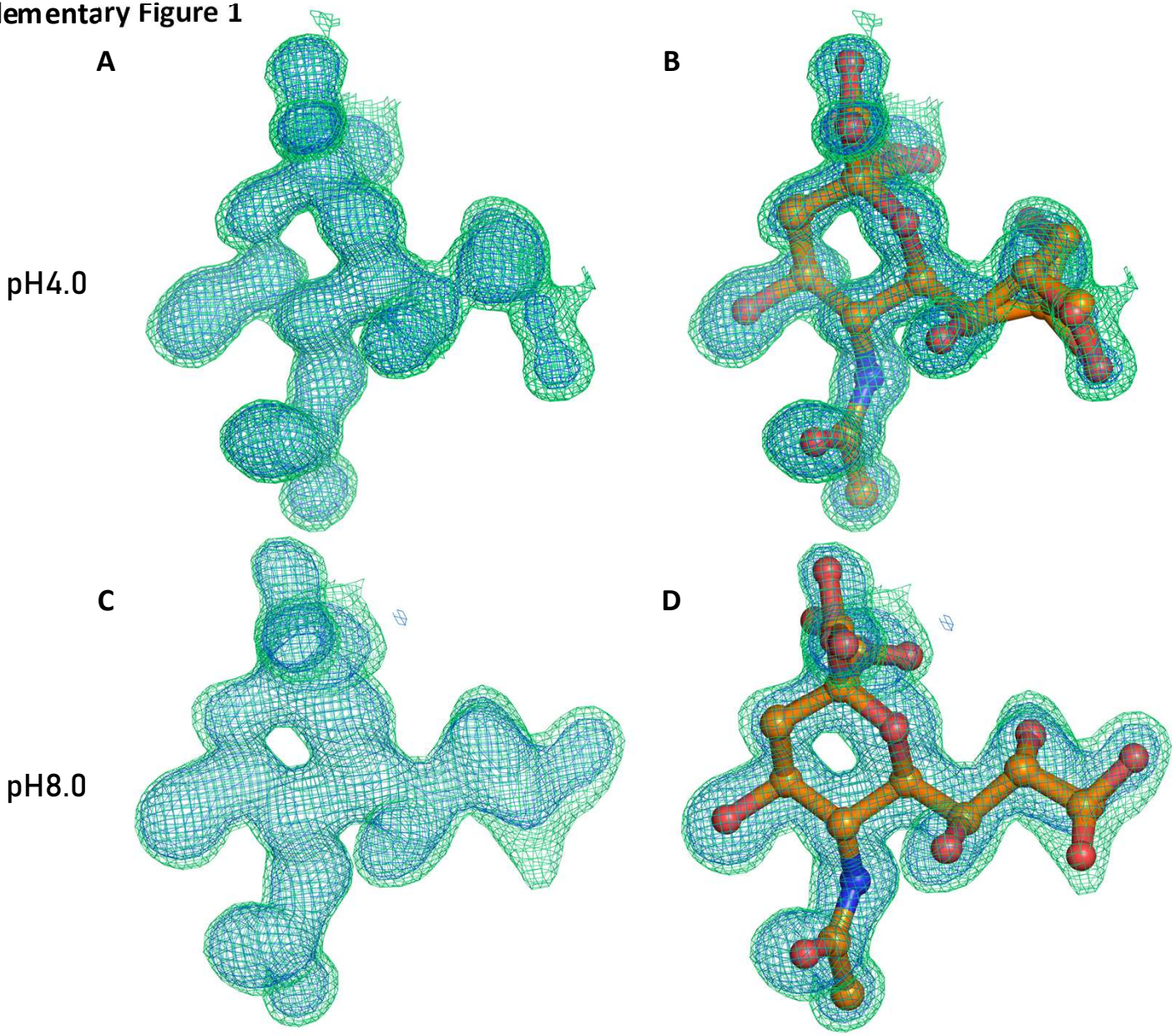

A
Lys349
Tyr320
Gln348
Tyr314
Thr319
Asn312
HAdV-D26
C
Ile310
Ile324
B
Lys345
Pro317
Tyr312
Ser344
Thr310
HAdV-D37
D
Tyr308
Val322

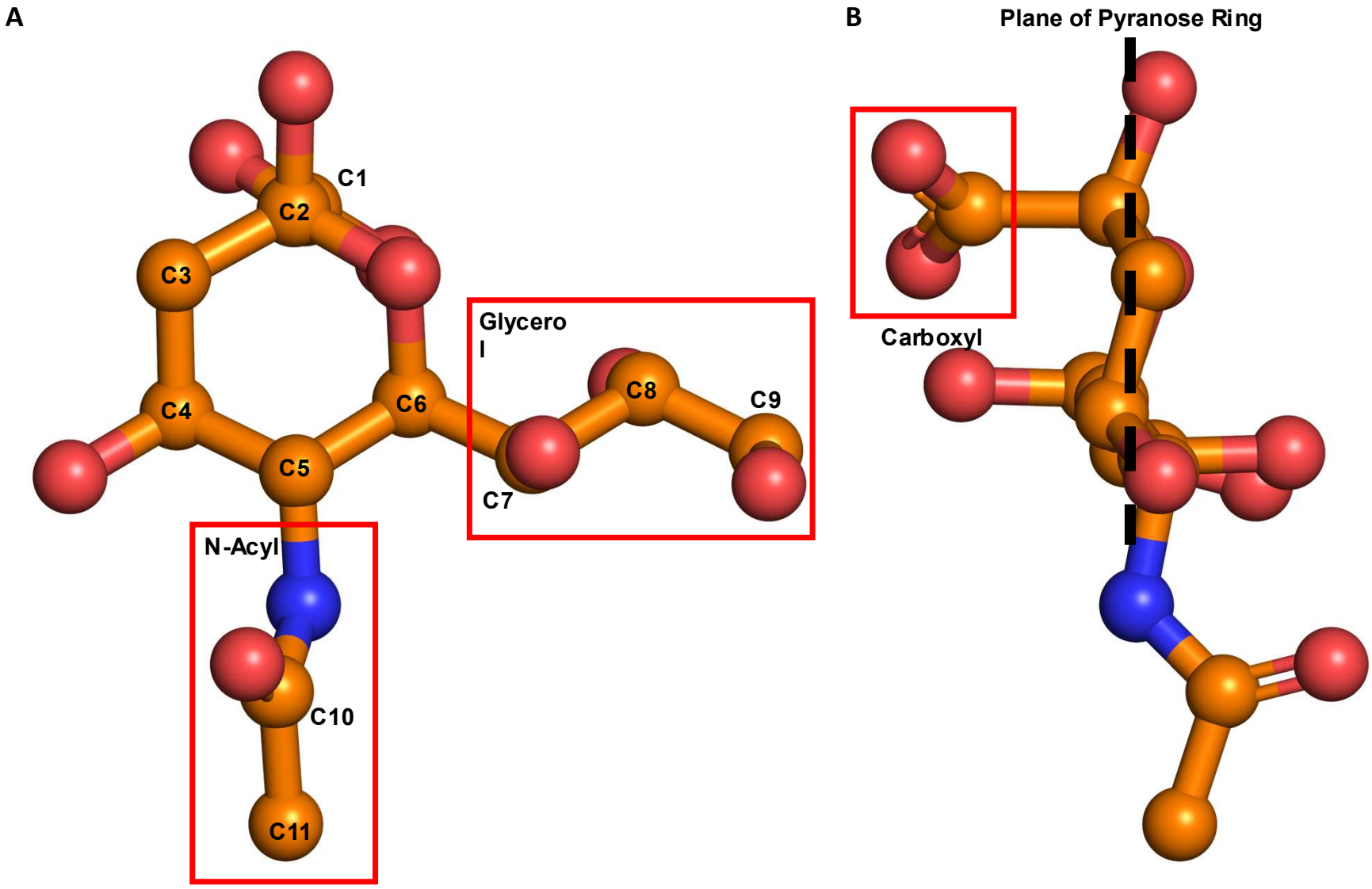
Figure 5


A
Gln348 – B
Gln348 – A
B
Gln348 – B
Gln348 – A
C
Tyr312
Ile324
Ile310
D
Tyr312
Ile324
Ile310

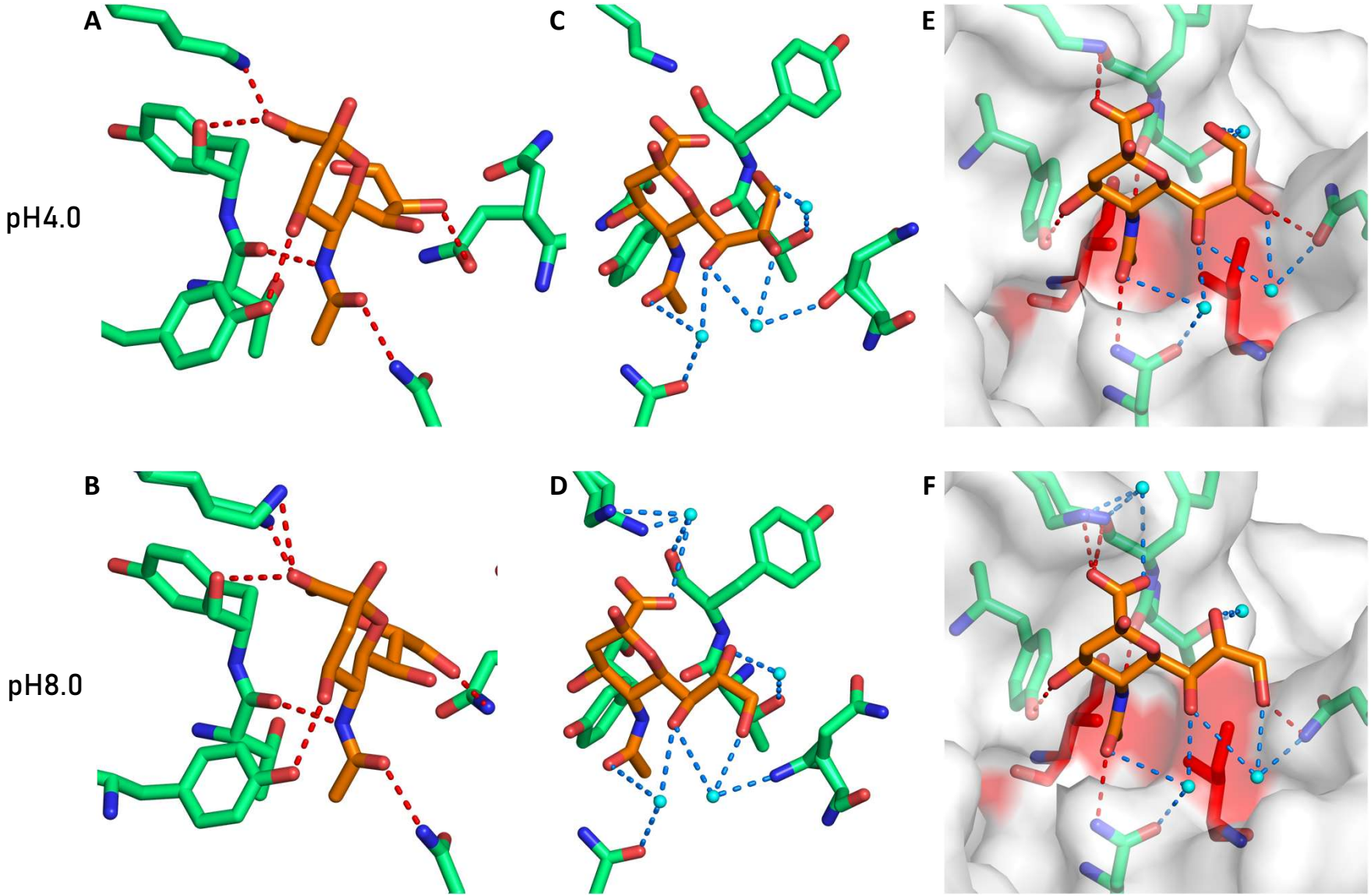
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4

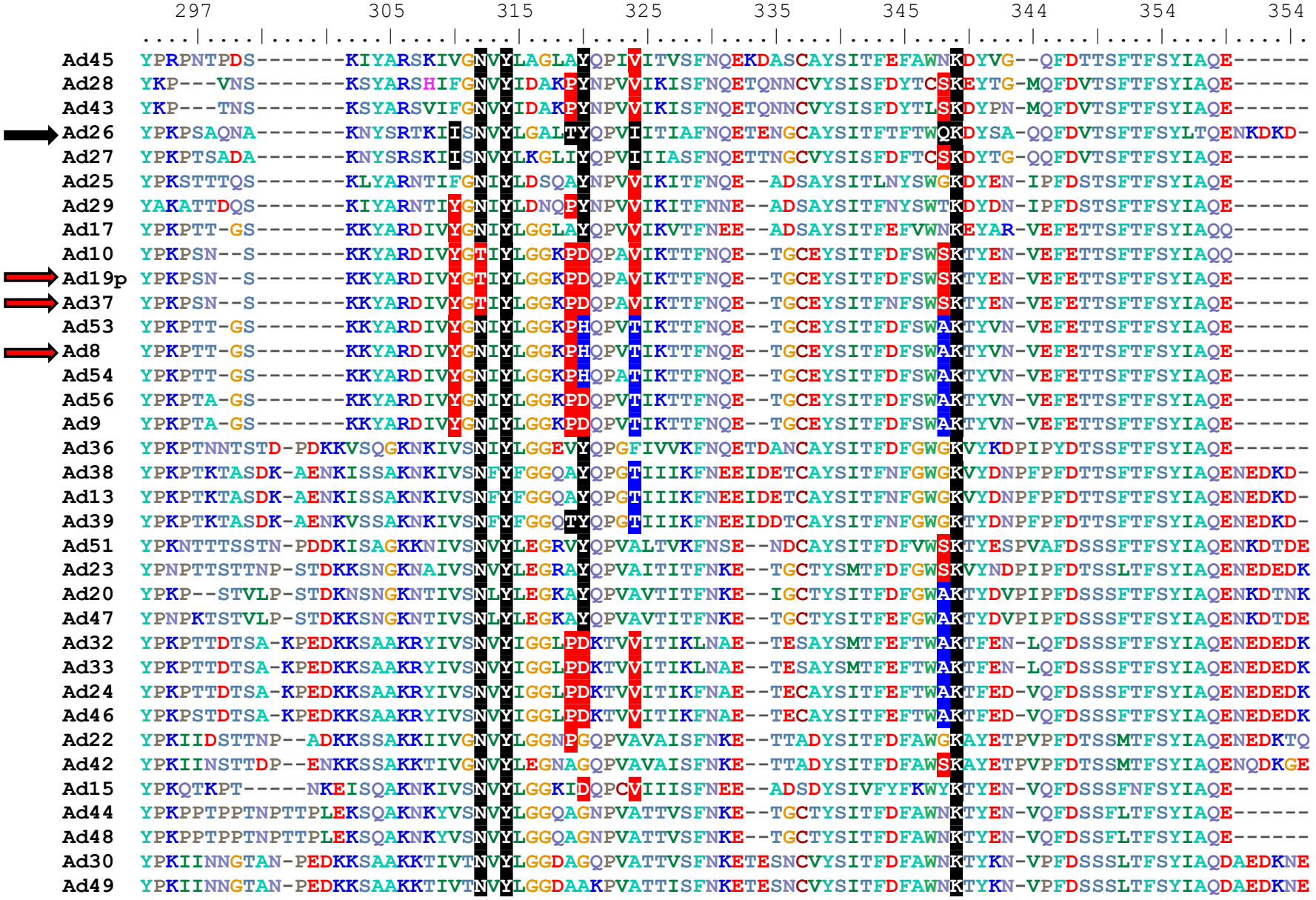


Table 1. Data collection and refinement statistics for structures generated in this study

PDB Entry	**6QU6**	**6QU8**	**6FJO**
Data Collection			
Diamond Beamline	I04	I04	I04
Date	25/10/2018	25/10/2018	05/12/2017
Wavelength	0.91587	0.91587	0.9795
Crystal Data			
Crystallisation Conditions	0.1 M MIB, 25 % w/v PEG 1500	0.1 M MIB, 25 % w/v PEG 1500	0.1 M SPG, 25% PEG 1500
pH	4.0	8.0	4.0
a=*b*=*c* (Å)	85.73	85.92	85.78
α=β=γ (°)	90.0	90.0	90.0
Space group	P 2_1 3	P 2_1 3	P 21 3
Resolution (Å)	1.03 – 49.5	1.19 – 42.96	1.17-85.78
Outer shell	1.03 – 1.06	1.19 – 1.22	1.17-1.23
R-merge (%)	5.3 (125.3)	8.5 (276.1)	6.5 (137.0)
R-meas (%)	5.5 (161.3)	8.8 (283.2)	6.6 (140.4)
CC1/2	1.0 (0.224)	1.0 (0.505)	1.00 (0.825)
I / σ(I)	21.9 (0.7)	20.7 (1.3)	24.8 (2.3)
Completeness (%)	97.7 (76.5)	100.0 (100.0)	100.0 (100.0)
Multiplicity	14.9 (2.0)	21.4 (20.3)	21.8 (21.2)
Total Measurements	1,509,159	1,448,478	1,568,641
Unique Reflections	103,975	67,799	71,878
Wilson B-factor(Å^2)	8.4	11.1	12.2
Refinement Statistics			
Total number of refined	1,936	1,846	1,761
R-work reflections	96,027	64,286	68,283
R-free reflections	4,907	3,441	3,558
R-work/R-free (%)	13.6 / 14.8	14.17 / 17.20	17.0 / 19.0
rms deviations			
Bond lengths (Å)	0.012	0.011	0.021
Bond Angles (°)	1.754	1.661	2.080
[1]Coordinate error	NULL	NULL	0.026
Mean B value (Å^2)	17.6	29.6	19.9
Ramachandran Statistics			
Favoured/allowed/Outliers	119 / 9 / 0	126 / 10 / 0	133 / 10 / 1
%	93.0 / 7.0 / 0.0	92.7 / 7.4 / 0.0	92.4 / 6.9 / 0.7

* One crystal was used for determining each structure.
* Figures in brackets refer to outer resolution shell, where applicable.
[1] Coordinate Estimated Standard Uncertainty in (Å), calculated based on maximum likelihood statistics.

Buffers:
- MIB: Malonic acid, Imidazole, Boric acid
- SPG: Succinic acid, Sodium phosphate monobasic monohydrate, Glycine: pH 4.0