



Structural and functional insights into species D adenovirus receptor usage: Implications for oncolytic virotherapy

Alexander T. Baker

Thesis submitted for the degree of Doctor of Philosophy



September 2019

Division of Cancer and Genetics
School of Medicine, Cardiff University
Cardiff, Wales, UK

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed (candidate)

Date

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed (candidate)

Date

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed (candidate)

Date

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed (candidate)

Date

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed (candidate)

Date

Acknowledgments

Thank you to Tenovus, and Cardiff University School of Medicine for funding and supporting me, and to the British Society for Gene and Cell Therapy for letting me be a part of such a brilliant society and supporting me as an ECR. Thank you to John Chester and Gavin Wilkinson for insightful conversations and words of genuine wisdom.

Thank you to my supervisor, Alan Parker, for having patience with my infamous tangents and giving me the rope to explore my ideas, but never enough to hang myself with. It is easy to become disillusioned by the area we work in. It has made all the difference to have a group leader who exemplifies the best of our field, not only as a scientist, but in your belief in and support for your team.

Thank you to the whole Parker Group and virology team, past and present. You all made me want to come to the lab each day. To Hanni, James, and Sarah, who taught me everything in the beginning and made it OK to laugh at knob jokes, you will always have my gratitude. Thank you to Evi and Simone for making sure I don't destroy expensive equipment, making my flow experiments work, and many late-night chats. To Pierre Rizkallah, for encouraging me into a field I was never supposed to work in and for making it a joy, I am sincerely grateful.

Perhaps most of all, thanks to all those who have believed in and supported me throughout this PhD. There are too many of you to ever list. Huge thanks are owed to those who have contributed to the scientific work I have done over these last (nearly) four years. You have all helped to make this thesis a reality and to make me a better scientist. Maybe more thanks are owed to those who have contributed to keeping me sane during the course of this work. Without your help I could not have finished.

The greatest thanks, however, are owed to those who have done both. To those who have helped me complete my work when it was hard, encouraged me when I have failed, listened to and helped me with everything I have been unsure about, both personal and scientific. It is an understatement to say I am eternally grateful.

Finally, thank you to the person who named it the Fiber-*knob* domain. I hope you have some idea of how much joy it created. It will never stop being funny.

Abstract

Adenoviruses are a diverse virus family, infecting a range of animal hosts. The human adenoviruses comprise over 100 types, divided into seven species, A-G. Adenoviruses are clinically important as both engineered therapeutic vectors and pathogens.

Adenoviruses are non-membrane bound viruses with double stranded DNA genomes, making them amenable to genetic manipulation and manufacturing at scale. This makes them attractive candidates as therapeutic vectors, currently under development as gene therapy vectors, viral vaccines, and oncolytic viruses. Adenovirus based vectors are in clinical trials for the prevention and treatment of diseases as diverse as Ebola infection and cancer, but are hindered by pre-existing antiviral immunity, leading to neutralisation of the virotherapy before it can have its therapeutic effects, and off-target tissue infections resulting in reduced delivery of therapeutic vector to the site of need.

Many natural adenovirus infections do not require medical intervention as they are easily neutralised by the hosts immune system, though they can result in serious disease. Predominantly, this is a concern for immunocompromised patients. However, some types can cause symptoms including gastroenteritis, epidemic keratoconjunctivitis (EKC), and potentially fatal acute respiratory distress syndrome, in healthy adults.

Species D adenoviruses are especially diverse, accounting for more than 50% of total adenovirus diversity and many instances of disease, including outbreaks of EKC. They can also be attractive as a basis for therapeutic viruses due to low rates of seroprevalence in the population and favourable immunogenicity profiles. However, little is known about most members of this species or the mechanisms which can make them either pathogenic or therapeutically useful.

The work presented herein seeks to better understand how species D adenoviruses infect cells. We therefore investigate their fiber proteins, which mediate primary cell surface receptor binding. We identify previously unknown adenovirus receptor interactions and examine these results in the context of developing new therapeutic adenovirus vectors.

Table of Contents

I	Declaration.....	II
II	Acknowledgments.....	III
III	Abstract	IV
IV	Table of Contents	V
V	List of Tables	XI
VI	List of Figures	XII
CHAPTER 1.	Introduction.....	1
1.1	Introduction.....	2
1.2	The Principle of Oncolytic Virotherapy.....	2
1.3	Adenovirus Biology.....	4
1.3.1	Adenovirus species.....	4
1.3.2	Neutralising antibodies in adenovirus infection	5
1.3.3	Adenovirus capsid structure	7
1.4	Disrupting natural adenovirus tropisms to reduce off-target infection	9
1.4.1	CAR (Coxsackie and Adenovirus Receptor)	12
1.4.2	CD46/MCP (Membrane Cofactor Protein).....	13
1.4.3	Desmoglein 2	17
1.4.4	Sialic acid, Glycan interaction, and GD1a	19
1.4.5	Polysialic acid.....	20
1.4.6	Sialic acid binding in the species D adenoviruses.....	21
1.4.7	Coagulation Factor X (FX), Heparan sulfate proteoglycans (HSPGs), and IgM.....	24
1.4.8	Integrins $\alpha\beta3$ and $\alpha\beta5$	27
1.4.9	MARCO/SR-A6 Scavenger Receptor (Macrophage Receptor with Collagenous Structure).....	29

1.5	Retargeting adenovirus vectors using engineered receptor proteins	29
1.5.1	Single chain antibodies (scFv)	30
1.5.2	Affibodies and FGFR2	33
1.5.3	Designed ankyrin repeat proteins (DARPs)	35
1.5.4	scTCR (single chain T-Cell Receptor) chimeric fiber proteins	36
1.5.5	Retargeting by peptide incorporation	37
1.5.6	The A20 peptide to target $\alpha v \beta 6$	37
1.6	Making adenovirus replication conditional upon oncogenic mutations in the cellular environment.....	40
1.7	Clinical trials of adenovirus derived oncolytic viruses.....	43
1.8	Aims	50
CHAPTER 2. Materials & Methods		51
2.1	Sequence comparison and Phylogenetic Tree generation	52
2.1.1	Sequence alignment and visualisation	52
2.1.2	Wu-Kabat variability analysis.....	52
2.1.3	Phylogenetic Maximum likelihood tree generation.....	52
2.2	Expression of recombinant fiber knob proteins	53
2.3	Cell Culture Methods.....	54
2.3.1	Culture of cell lines	54
2.3.2	Long term cell storage	56
2.3.3	Monocyte derived dendritic cells (MDDC).....	57
2.4	Virus generation	57
2.4.1	Generation of pseudotype virus bacterial artificial chromosomes....	57
2.4.2	Amplification of BAC DNA	59
2.4.3	Transfection of BAC DNA and virus rescue.....	60
2.4.4	Propagation of rescued viruses	61

2.4.5	Purification of viruses into high titre stocks.....	61
2.4.6	Purification of viruses into high titre stocks.....	62
2.5	Infectivity Assays	63
2.5.1	GFP infectivity assay.....	63
2.5.2	Luciferase infectivity assay.....	64
2.5.3	Receptor blocking assay	64
2.5.4	Neuraminidase and Heparinase cleavage assays	65
2.5.5	Competition inhibition assay.....	66
2.5.6	Statistical analysis of infectivity assays	66
2.6	Structural Analysis of fiber knob proteins.....	67
2.6.1	Homology modelling of HAdV-D8K	67
2.6.2	Crystallisation of fiber knob proteins.....	67
2.6.3	Soaking with sialic acid	67
2.6.4	X-Ray structure data collection	67
2.6.5	Calculation of electrostatic surface profile	68
2.6.6	Fiber knob to receptor docking.....	68
2.6.7	Calculation of structure interface energies	68
2.7	Stable Isotope Labelling of Amino acids in Cells (SILAC).....	69
2.8	Determination of receptor binding affinity by surface plasmon resonance (SPR)	70
CHAPTER 3. Results: Investigating the broad tropism of species D adenovirus type 49		
		72
3.1	Introduction:	73
3.2	Phylogenetic analysis of the fiber knob domain reveals hidden diversity within the species D	76
3.3	Adenovirus fiber knob pseudotypes exhibit different infectivity characteristics in test cell lines	78

3.4	The HAdV-C5/D49K pseudotyped virus infects cancer cell lines with high efficiency.....	87
3.5	HAdV-C5/D49K can infect dendritic cells with high efficiency	92
3.6	The structure of the HAdV-D49 fiber knob domain is highly homologous to that of HAdV-D30 with a different infectivity profile	97
3.7	SILAC Co-precipitation suggests that HAdV-D49K may interact with an adhesion protein	106
3.8	Discussion.....	112
3.8.1	Adenovirus diversity is under-represented by whole virus phylogenetics	113
3.8.2	Infectivity studies using HAdV-C5 pseudotypes suggest previously unknown receptor interactions within the species D adenoviruses	114
3.8.3	Structural comparison reveals electrostatic surface profile is key to the broad tropism of HAdV-D49K.....	116
3.8.4	Pull down implicates the desmosome in HAdV-D49K mediated infection.....	117
3.8.5	Infectivity studies of HAdV-C5/D49K in cancer cell lines reveal high efficiency infection and pan-tropism.....	120
3.8.6	HAdV-C5/D49K can infect dendritic cells with high efficiency and may have utility as a vaccine vector.....	121
3.9	Summary of Chapter 3.....	124
CHAPTER 4. Results: Characterisation of three species D adenovirus fiber knob proteins and their interaction with known adenovirus receptors		
4.1	Introduction:	126
4.2	Novel species D fiber knob crystal structures:.....	128
4.2.1	Expression and purification of recombinant fiber knob proteins.....	128
4.2.2	Novel crystal structures of human adenovirus fiber knob proteins:	130
4.2.3	Biological trimers determined from crystallographic structures:.....	137

4.3 Examination of species D adenovirus fiber knob binding to known adenovirus receptor proteins:.....	138
4.3.1 Investigated species D adenovirus fiber knobs are unlikely to utilise CD46 as a primary receptor:.....	138
4.3.2 Desmoglein 2 is unlikely to be the primary receptor for the investigated species D adenovirus fiber knobs:.....	144
4.3.3 Adenoviruses D26, D48, and D49 fiber knob proteins bind CAR with low affinity and suggest regulation through steric hindrance:	145
4.4 Discussion:	154
4.5 Summary of Chapter 4.....	160
 CHAPTER 5. Results: Evaluation of Species D adenovirus fiber knob tropisms	
161	
5.1 Introduction:.....	162
5.2 HAdV-D26 fiber knob has homology with known sialic acid utilising adenoviruses.....	164
5.3 HAdV-D26 pseudotyped viruses require sialic acid for efficient cell entry	166
5.4 Crystal structures reveal sialic acid binding in the apex of HAdV-D26K	168
5.4.1 Sialic acid binding is stable at high and low pH.....	168
5.4.2 HAdV-D26K binding to sialic acid utilises multiple types of bond..	172
5.5 HAdV-D26 binds to sialic acid in a manner similar to known sialic acid utilising species D adenoviruses	177
5.6 Species D adenoviruses conserve sialic acid binding features.....	180
5.7 Discussion	195
5.7.1 pH is unlikely to affect sialic acid binding to HAdV-D26K.....	195
5.7.2 HAdV-D26K contains residue substitutions in the sialic acid binding pocket enabling more direct bonds	196

5.7.3	Regions of weak density around residues with long side chains may indicate an induced fit model of sialic acid binding.....	197
5.7.4	HAdV-D26K and HAdV-D19pK may belong to a wider subset of non-EKC causing sialic acid binding adenoviruses	198
5.7.5	Fiber knob binding to sialic acid does not preclude the possibility of binding to alternative sugar monomers	199
5.7.6	Neuraminidase assays suggest many species D adenovirus fiber knobs can utilise sialic acid as a receptor to varying degrees	201
5.8	Summary of chapter 5.....	204
CHAPTER 6. General Discussion		205
6.1	Adenovirus receptor tropism as a driver of viral evolution.....	207
6.2	Further investigation of species D adenovirus receptor tropism is imperative to understanding its pathogenesis.....	210
6.3	The potential for species D adenoviruses as therapeutics	211
6.3.1	Towards target specific species D adenovirus vectors	212
6.3.2	The potential of species D adenoviruses in vaccine development.....	213
6.4	Concluding remarks	214
VII	References	XVII
VIII	Publications	LXXIII

List of Tables

TABLE 1-1 ADENOVIRUS RECEPTORS, THE RESIDUES SHOWN TO FACILITATE RECEPTOR INTERACTIONS, AND DEMONSTRATED TROPISM ABLATING MUTATIONS.....	10
TABLE 1-2: A SUBSET OF ACTIVE CLINICAL TRIALS UTILISING ADENOVIRUS DERIVED ONCOLYTIC VIRUSES.....	45
TABLE 2-1 CELL LINES USED IN THIS STUDY, THEIR ORIGINS, AND THE MEDIA WHICH WAS USED TO SUPPORT THEM.....	54
TABLE 3-1 DATA COLLECTION AND REFINEMENT STATISTICS FOR HADV-D49K (6QPN) AND HADV-D30K (6QPM).	99
TABLE 4-1 STATISTICS OF THE ADENOVIRUS FIBER-KNOB CRYSTAL STRUCTURES GENERATED IN THIS STUDY, AS DEPOSITED IN THE WWPDB.....	132
TABLE 4-2 SURFACE PLASMON RESONANCE (SPR) MEASUREMENTS CONFIRM FINDING REGARDING CD46 AND CAR AFFINITY AND RULE OUT DESMOGLEIN 2 AS A POTENTIAL RECEPTOR FOR THESE SPECIES D ADENOVIRUSES.	144
TABLE 5-1 REFINEMENT STATISTICS FOR STRUCTURES OF HADV-D26K USED IN THIS STUDY.....	168
TABLE 5-2 REFINEMENT STATISTICS FOR DEPOSITED STRUCTURES OF PREVIOUSLY UNDESCRIBED SPECIES D ADENOVIRUSES.	181

List of Figures

FIGURE 1-1 CARTOON VIEW OF ADENOVIRUS HIGHLIGHTING THE MAJOR CAPSID PROTEINS, LABELLED.	7
FIGURE 1-2 CAR INTERACTING RESIDUES WITHIN THE HADV-C5 FIBER KNOB DOMAIN.....	12
FIGURE 1-3 CD46 INTERACTING RESIDUES, AND KNOWN MUTATION SITES, WITHIN THE HADV-B35 FIBER KNOB DOMAIN:	16
FIGURE 1-4 DSG2 INTERACTING RESIDUES, AND KNOWN MUTATION SITES, WITHIN THE HADV-B3 FIBER KNOB DOMAIN.	18
FIGURE 1-5 GD1A/SIALIC ACID INTERACTING RESIDUES WITHIN THE HADV-D37 FIBER KNOB DOMAIN.	23
FIGURE 1-6 OVERVIEW OF ANTIBODY-LIKE PROTEINS SPECIFIC TO HER2 WHICH HAVE BEEN GENETICALLY INTEGRATED INTO ADENOVIRUS.	32
FIGURE 1-7 MAP OF THE NUMBER AND LOCATIONS OF ACTIVE CLINICAL TRIALS USING ADENOVIRUS VECTORS FROM CLINICALTRIALS.GOV USING THE FOLLOWING SEARCH CRITERIA: ON 30/07/2019:	48
FIGURE 3-1 PHYLOGENETIC ANALYSIS OF ADENOVIRUS GENOMES AND FIBER KNOB DOMAINS REVEALS INCREASED DIVERSITY IN THE SPECIES D ADENOVIRUSES.....	77
FIGURE 3-2 TRANSDUCTION EFFICIENCY OF ADENOVIRUS FIBER PSEUDOTYPES IN CHO CELL LINES EXPRESSING KNOWN PRIMARY ADENOVIRUS RECEPTORS, OR NO KNOWN ADENOVIRUS RECEPTORS.....	80
FIGURE 3-3 REMOVAL OF CELL SURFACE HSPGS BY HEPARINASE TREATMENT DOES NOT AFFECT HADV-C5/D49K TRANSDUCTION EFFICIENCY.....	83
FIGURE 3-4 REMOVAL OF CELL SURFACE SIALIC ACID BY NEURAMINIDASE TREATMENT DOES NOT DIMINISH HADV-C5/D49K TRANSDUCTION EFFICIENCY.	84
FIGURE 3-5 HADV-C5 TRANSDUCTION IN CHO-CAR CELLS IS INHIBITED BY HADV-C5K AND HADV-D49K, BUT HADV-C5/D49K IS NOT INHIBITED	

BY EITHER, SUGGESTING A CAR INDEPENDENT INFECTION PATHWAY.....	85
FIGURE 3-6 HADV-C5 TRANSDUCTION IN CHO-K1 CELLS IS INEFFICIENT AND NOT SIGNIFICANTLY AFFECTED BY PHARMACOLOGICAL BLOCKADE, WHILE HADV-C5/D49K TRANSDUCTION CAN BE INHIBITED BY BLOCKADE WITH ITS OWN FIBER KNOB PROTEIN.	86
FIGURE 3-7 HADV-C5/D49K CAN TRANSDUCE HUMAN PANCREATIC CANCER CELL LINES WITH HIGH EFFICIENCY. TRANSDUCTION WAS MORE EFFICIENT WITH HADV-C5/D49K THAN WITH HADV-C5 IN ALL TESTED PANCREATIC CANCER CELL LINES (NAMES GIVEN ABOVE PLOTS).....	89
FIGURE 3-8 HADV-C5/D49K CAN TRANSDUCE HUMAN BREAST CANCER CELL LINES WITH HIGH EFFICIENCY. TRANSDUCTION WAS MORE EFFICIENT WITH HADV-C5/D49K THAN WITH HADV-C5 IN ALL TESTED BREAST CANCER CELL LINES (NAMES GIVEN ABOVE PLOTS).....	90
FIGURE 3-9 HADV-C5/D49K CAN TRANSDUCE HUMAN LUNG, OESOPHAGEAL, COLORECTAL, AND OVARIAN CANCER CELL LINES WITH HIGH EFFICIENCY.	91
FIGURE 3-10 TRANSFECTION OF DONOR HUMAN MONOCYTE DERIVED DENDRITIC CELLS (HMDDC) IS MORE EFFICIENT BY HADV-C5/D49K THAN BY HADV-C5.	94
FIGURE 3-11 TRANSDUCTION OF DONOR HMDDC IS MORE EFFICIENT USING HADV-C5/D49K, BUT MAY RESULT IN INCREASED CELL DEATH.	95
FIGURE 3-12 TRANSDUCTION OF MOUSE MDDC SHOWS SIMILAR IMPROVEMENT IN EFFICIENCY USING HADV-C5/D49K RATHER THAN HADV-C5.....	96
FIGURE 3-13 HADV-C5/D49K AND HADV-C5/D30K SHOW DIFFERENT INFECTIVITY PROFILES DESPITE DIFFERING IN ONLY FOUR RESIDUES.	98
FIGURE 3-14 OVERVIEW OF THE SOLVED CRYSTAL STRUCTURE OF HADV-D30K, DEPOSITED IN THE WWPDB AS 6STU.	101

FIGURE 3-15 OVERVIEW OF THE SOLVED CRYSTAL STRUCTURE OF HADV-D49K, DEPOSITED IN THE WWPDB AS 6QPN.....	102
FIGURE 3-16 NON-CONSERVED RESIDUES IN HADV-D30K AND HADV-D49K ARE SIMILARLY ORIENTED IN SPACE AT THE APEX OF THE FIBER KNOB.....	103
FIGURE 3-17 THE DIFFERENCES IN RESIDUES BETWEEN HADV-D30K AND HADV-D49K HAVE A LARGE EFFECT ON THE SURFACE ELECTROSTATIC POTENTIAL OF THE FIBER KNOB.....	105
FIGURE 3-18 PROTEINS ENRICHED UPON AFFINITY PULL DOWN ASSAY PERFORMED WITH RECOMBINANT RAD49KN PROTEIN AS DETERMINED BY SILAC MASS SPECTROMETRY IN LABELLED SKOV-3 CELLS.....	110
FIGURE 4-1 EXAMPLE OF PURIFICATION OF FIBER-KNOB PROTEIN BY NI-IMAC AND SEC.....	129
FIGURE 4-2 OVERVIEW OF THE SOLVED CRYSTAL STRUCTURE OF HADV-D26K, DEPOSITED IN THE WWPDB AS 6FJN.....	133
FIGURE 4-3 OVERVIEW OF THE SOLVED CRYSTAL STRUCTURE OF HADV-D48K, DEPOSITED IN THE WWPDB AS 6FJQ.....	134
FIGURE 4-4 OVERVIEW OF THE SOLVED CRYSTAL STRUCTURE OF HADV-C5K, DEPOSITED IN THE WWPDB AS 6HCN.....	135
FIGURE 4-5 OVERVIEW OF NOVEL CRYSTAL STRUCTURES OF SPECIES C AND D ADENOVIRUSES FIBER-KNOB TRIMERS GENERATED IN THIS STUDY.....	136
FIGURE 4-6 STABILITY OF ADENOVIRUS FIBER-KNOB TRIMERS.....	137
FIGURE 4-7 SPECIES D ADENOVIRUS FIBER-KNOBS ARE UNLIKELY TO FORM A STABLE COMPLEX WITH CD46.....	141
FIGURE 4-8 COMPETITION INHIBITION ASSAYS SUGGEST SPECIES D ADENOVIRUS FIBER-KNOBS CANNOT BIND CD46.....	143
FIGURE 4-9 ADENOVIRUS FIBER-KNOB INTERACTION WITH COXSACKIE AND ADENOVIRUS RECEPTOR (CAR) IS MADE UP OF TWO INTERFACES, AND STERIC HINDERANCE BY THE DG LOOP OF SPECIES D ADENOVIRUSES MAY INHIBIT CAR INTERACTION.....	148

FIGURE 4-10 COMPETITION INHIBITION ASSAYS SUGGEST SPECIES D ADENOVIRUS FIBER-KNOBS BIND TO CAR WITH VARIABLE AFFINITY.	150
FIGURE 4-11 EFFICIENCY OF INFECTION BY ADENOVIRUS FIBER-KNOB PSEUDOTYPES IS PROPORTIONAL TO INCUBATION TIME IN CHO-CAR CELLS.	152
FIGURE 4-12 THE PROLINE RICH LOOP OF HADV-D48K CONTRIBUTES TO A LOW CAR AFFINITY.....	153
FIGURE 5-1 HADV-D26K SHARES KEY SIALIC ACID BINDING RESIDUES WITH EKC ASSOCIATED ADENOVIRUSES (A) BUT HAS LOW OVERALL SEQUENCE IDENTITY AND A VERY DIFFERENT PREDICTED ISOELECTRIC POINT (B).	164
FIGURE 5-2 COMPARISON OF THE ELECTROSTATIC SURFACE POTENTIAL OF EKC ASSOCIATED ADENOVIRUSES.....	165
FIGURE 5-3 NEURAMINIDASE TREATMENT SIGNIFICANTLY INHIBITS INFECTION BY HADV-C5/D26K, BUT NOT HADV-C5 OR HADV-C5/B35K, IN THREE CANCER CELL LINES.....	167
FIGURE 5-4 CRYSTAL STRUCTURES SHOW SIALIC ACID INTERACTS WITH HADV-D26K IN THE APICAL DEPRESSION AT HIGH AND LOW PH...	170
FIGURE 5-5 THE BIOLOGICALLY RELEVANT CONFORMATION OF SIALIC ACID, WHEN IN COMPLEX WITH HADV-D26K.....	171
FIGURE 5-6 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.....	175
FIGURE 5-7 HADV-D26K AFFECTS AN INDUCED FIT MECHANISM IN SIALIC ACID BINDING.	176
FIGURE 5-8 HADV-D26K FORMS A COMPLEX INTERACTION NETWORK OF HYDROPHOBIC AND ELECTROSTATIC INTERACTIONS WITH SIALIC ACID.....	179
FIGURE 5-9 SPECIES D ADENOVIRUSES CONSERVE KNOWN SIALIC ACID BINDING RESIDUES.	182

FIGURE 5-10 OVERVIEW OF THE SOLVED CRYSTAL STRUCTURE OF HADV-D10K, DEPOSITED IN THE WWPDB AS 6QPM.....	184
FIGURE 5-11 OVERVIEW OF THE SOLVED CRYSTAL STRUCTURE OF HADV-D15K, DEPOSITED IN THE WWPDB AS 6STW.....	185
FIGURE 5-12 OVERVIEW OF THE SOLVED CRYSTAL STRUCTURE OF HADV-D29K, DEPOSITED IN THE WWPDB AS 6STV.....	186
FIGURE 5-13 SIALIC ACID BINDING RESIDUES OF HADV-D26K AND THE EQUIVALENT RESIDUES, AS DEFINED BY AMINO ACID SEQUENCE ALIGNMENT, IN OTHER SPECIES D ADENOVIRUS CRYSTAL STRUCTURES.....	187
FIGURE 5-14 DATA COLLECTED FROM SMALL MOLECULE SOAKED HADV- D29K CRYSTALS SHOWS DENSITY CORRESPONDING TO SIALIC ACID BINDING WHILE SIMILAR DATA FROM HADV-D30K SHOWS NON- DESCRIPT SMALL MOLECULE BINDING.....	191
FIGURE 5-15 SOAKED STRUCTURE OF HADV-D15K DOES NOT SHOW CLEAR SIALIC ACID BINDING DENSITY, AND HAS A VERY TIGHT CRYSTAL PACKING ARRANGEMENT.....	192
FIGURE 5-16 SIALIC ACID CLEAVAGE SIGNIFICANTLY REDUCES THE INFECTIVITY OF 5 OF 6 TESTED HADV-C5 VIRUSES PSEUDOTYPED WITH SPECIES D FIBER-KNOB DOMAINS.	194

CHAPTER 1. Introduction

1.1 Introduction

The work in this thesis focuses on the virology of adenoviruses and how this understanding can be exploited in the design of new adenovirus vectors which overcome limitations related to targeted administration of virotherapies to cancer cells. As such, this introduction will focus upon the engineering of adenoviruses to specifically infect cancer cells, rather than healthy tissues. Another major topic within oncolytic virotherapy is utilising the viruses to manipulate the immune system to instigate an anti-tumour immune response.

Viruses are inherently immunogenic as they are recognised by the immune system as foreign agents, leading to clearance, this can be beneficial in the context of virotherapies by recruiting T-cells to the tumours and breaking immunological tolerance. There are numerous virotherapies expressing immunomodulatory transgenes in attempts to better stimulate anti-tumour immunity. This review will discuss adenovirus and immunomodulation in the context of manipulating the native adenovirus biology, rather than adding transgenes or manipulating replication, which are post-infection enhancers of oncolysis rather than a means of viral targeting, and therefore beyond the scope.

1.2 The Principle of Oncolytic Virotherapy

Viruses were once described by Peter Medawar, the 1960 Nobel prize recipient for the discovery of immunological tolerance, as “a piece of bad news wrapped up in protein”¹. Whilst this is an apt description of viruses in disease, it now stings of irony. Viruses have emerged as clinically important agents in the treatment of many diseases, including cancer, where one of its key modes of action is breaking of immune tolerance to tumour cells².

The idea that viruses can of benefit to cancer patients has existed since at least 1904, when it was observed that influenza infection can be beneficial for leukaemia patients³. However the effort to harness viruses as cancer therapeutics were stunted until the development of cell culture techniques, and clinical trials began in the 1950's⁴. Viruses were first reported to destroy sarcoma in mice in 1949, using Russian Far East Encephalitis virus⁵, and shown to specifically kill cancer cells as long ago as 1953, when Koprowski *et al*

stated that “*Studies on the specificity of such a reaction for a given virus[...] may facilitate an understanding of the mechanism of selective cell destruction*”⁶, predicting a now global research effort to harness viruses as anticancer therapeutics.

The discovery of DNA and ways to engineer it, as well as *ex vivo* cell culture and viral propagation techniques, have enabled the investigation of a wide range of different viruses as cancer therapeutics. These range from wildtype small RNA viruses like Reovirus (Respiratory Enteric Orphan virus), to engineered large DNA viruses such as Herpes Simplex^{7,8}.

The original principle underpinning the idea of oncolytic virology was to take advantage of the natural viral characteristics. These are a natural propensity to replicate more efficiently in rapidly dividing and dysregulated cells, such as cancer cells deficient in p53⁹, or innate mechanisms of defence such as the interferon pathway¹⁰. This remains core to many oncolytic virotherapy strategies, such as those mediated by unmodified reovirus, or vesicular stomatitis virus. It is now generally accepted that the viruses also serve to stimulate the immune system to attack the tumour and assist in clearance, making oncolytic viruses a form of immunotherapy¹¹.

Wild type viruses have been reported as effective interventions for cancer in isolated instances^{4,12,13}. However, they have numerous limitations and risks associated with them, including pathogenesis, dose limiting toxicity, weak immunogenicity or rapid clearance, and low efficiency of delivery to the tumour. This combines to make them unpredictable agents with a poor safety profile.

Wildtype viruses can be engineered to overcome these challenges. Broadly speaking, these engineering methods are divided into two categories: rational design and directed evolution¹⁴. Rational design is based upon an intricate understanding of the virus’s native pathways and making deliberate changes to these pathways to achieve the desired effects when administered as a drug. Directed evolution approaches allow natural recombination and mutation under controlled conditions in order to select for a virus with desirable traits.

1.3 Adenovirus Biology

1.3.1 Adenovirus species

The international committee on taxonomy of viruses (ICTV) defines 7 adenovirus species. However, they do not delineate the viruses any further, avoiding the contentious issue of how to define an individual adenovirus type¹⁵.

Canonically there are 57 adenovirus types divided into the 7 species, A-G. These types were determined by serological testing for neutralising antibody activity and haemagglutination¹⁶. This method of defining adenovirus serotypes is extremely coarse due to the potential for cross-neutralisation. This is typified by the confusion surrounding Human Adenovirus Species D type 19 (HAdV-D19). The prototype HAdV-D19 virus, HAdV-D19p, was first isolated in Saudi Arabia in 1955¹⁷ and is not a cause of Epidemic Keratoconjunctivitis (EKC). However, another virus, HAdV-D19a, is a cause of EKC and possesses a similar serological fingerprint¹⁸. It has now been shown to have a very different genome, helping to explain its different pathology¹⁹. This has led to HAdV-D19a, the EKC causative agent, being reclassified as HAdV-D64, leaving HAdV-D19p as the sole HAdV-D19. This shows how serological methods can misidentify an adenovirus and is neither surprising, given the ease with which they recombine to generate new recombinant types and alter the serological determinants^{20,21}, or the first time such misidentifications have occurred²²⁻²⁴.

Phylogenetic analysis of the species D adenoviruses (the most diverse of the human adenoviral species) reveals homologous recombination between hypervariable regions in each of the major capsid proteins and in every member of this subtype. In turn, there will be a direct impact on capsid structure and thus immune recognition²⁰. The authors of this study estimate that serology testing (the standard on which the canonical 57 serotypes is based) can only identify 53% of 38 fully characterised Species D adenoviruses unambiguously²⁰. An alternative taxonomic proposal attempts to reflect the genetic variation observed in the \approx 35Kbp linear dsDNA genomes of isolated adenoviruses, yielding 103 candidate serotypes as of July 2019^{25,26}.

This alternative taxonomic proposal is championed by the human adenovirus working group (HAdVWG). This method uses the genetic sequence of the three

major capsid proteins (described in the next section), the Penton, Hexon, and Fiber proteins, as the determinants of adenovirus type, defining a new type as any new combination of these three proteins. However, how to define a new type of these proteins is difficult as some viral proteins may differ in very few amino acids. Similar proteins may belong to very different adenovirus types with different clinical profiles, as is the case for the fiber proteins of HAdV-D37 and HAdV-D19p. While the fiber knob protein of HAdV-D37 (HAdV-D37K) is identical to that of HAdV-D64 (formerly HAdV-D19a) it differs with HAdV-D19pK by only 2 amino acids²⁷.

There have been proposals regarding the best ways to define these proteins based on cut-offs defining the degree of similarity²⁸. However, adenoviruses belonging to the same type may acquire mutations, potentially resulting in greater sequence differences than those between the HAdV-D37 and D19p fiber but may not cause any changes in clinical presentation. This makes the definition of a significant difference difficult. For now, no consensus has been reached on an answer to the typing problem. The solution for researchers reporting their findings is to work on carefully defined viruses and to report the viral protein sequences which their findings relate to alongside their results.

1.3.2 Neutralising antibodies in adenovirus infection

Neutralising antibodies (nAbs) are a key consideration in all adenovirus research. Firstly, they are the traditional means by which adenoviruses are typed. By utilising serum from patients infected with particular adenovirus types the way in which the adenoviruses relate to each other can be determined by the ability of different patient sera to neutralise the virus, rendering it unable to cause haemagglutination and/or infect cells²⁹. This is the source of the former name for adenovirus types: serotype. There are numerous variations on this test, each purporting to make serology based identification of adenoviruses more rapid and accurate²⁹⁻³¹. However, as mentioned in the previous section, this method of typing adenoviruses remains controversial. Serotyping provides no information about internal differences between viruses, as it is based upon nAb recognition of the solvent exposed capsid regions. Further, viruses may be neutralised by similar nAbs but have different structures as was observed with the incorrect typing of HAdV-D19a and subsequent retyping as HAdV-D64¹⁹.

Adenoviruses which have very similar proteins have also been shown to have different pathologies, as observed with the fiber proteins of HAdV-D37 and HAdV-D19p²⁷. Recombination's between viruses could therefore result in broadly similar capsids with different amino acids at key locations which alter their pathology. Such subtle recombination's have been observed repeatedly in the human adenoviruses, though their effect on pathogenicity is as yet unknown^{21,25,32,33}. To further support this, genetic analysis has already evidenced a novel adenovirus type, not identified through serology³⁴.

Neutralising antibodies are also a critical consideration in the development of adenovirus derived therapeutics, such as vaccines, gene transfer vectors, and oncolytic virotherapies. If a patient has nAbs against the serotype of adenovirus being utilised as a vector it may be neutralised before it is capable of having its therapeutic effects³⁵. This is further complicated by the fact that populations from different geographical areas have different levels of pre-existing immunity to different adenovirus serotypes. For example, an international study of patient epidemiology showed that 61-78% of patients had high or very high levels of HAdV-C5 nAbs, compared to 0.4-5.0% with nAbs against HAdV-B35. Within the Ad5 serotype there was a great deal of variance, approximately 27% of American adults at low risk of HIV infection had anti-HAdV-C5 nAbs, compared to 79% of high risk Kenyans³⁶. Some adenovirus species are generally lower seroprevalence than others though, with multiple studies showing species D adenoviruses having lower seroprevalence in all tested populations compared to the extremely common HAdV-C5 serotype³⁶⁻³⁹. This means that adenovirus-based therapeutics must be developed such that they either use rare adenovirus species which patients do not have pre-existing nAbs against, as with the selection of the HAdV-D26 virus as a vaccine vector³⁹⁻⁴¹, or engineered to be unrecognised by pre-existing nAbs, as with the anti-HIV vaccine, using adenovirus pseudotyped with the HAdV-D48 hexon variable regions^{42,43}. Epitope selection is also key to the design of antiviral vaccines, as key epitopes for viral function can be targeted for neutralisation without the virus being easily able to mutate to evade neutralisation, as has been achieved in adenovirus type 4 and 7^{44,45}.

1.3.3 Adenovirus capsid structure

Adenovirus genomes are approximately 35kbp long and made of double stranded DNA (dsDNA). The exact makeup of an adenovirus genome varies depending on type, and the specific isolate. However, the overall structure of the virus capsid remains broadly similar, consisting of an icosahedral capsid comprised of 3 major proteins, the penton base, the hexon, and the fiber (Fig1-1A), and 4 minor proteins (Fig1-1B)^{14,46}.

The hexon is the most abundant capsid protein and makes up the 20 facets of the icosahedral viral capsid. This gives a total of 120 hexon trimers (360 monomers) per viral particle. Each hexon is a homotrimer, which gives the hexon the eponymous hexagonal tertiary structure, yielding 360 monomers per viral capsid. The hexagonal base of the protein tessellates to form the capsid and projections extend away from the capsid. The projections are highly variable (the hypervariable regions, HVRs) between species and are a major

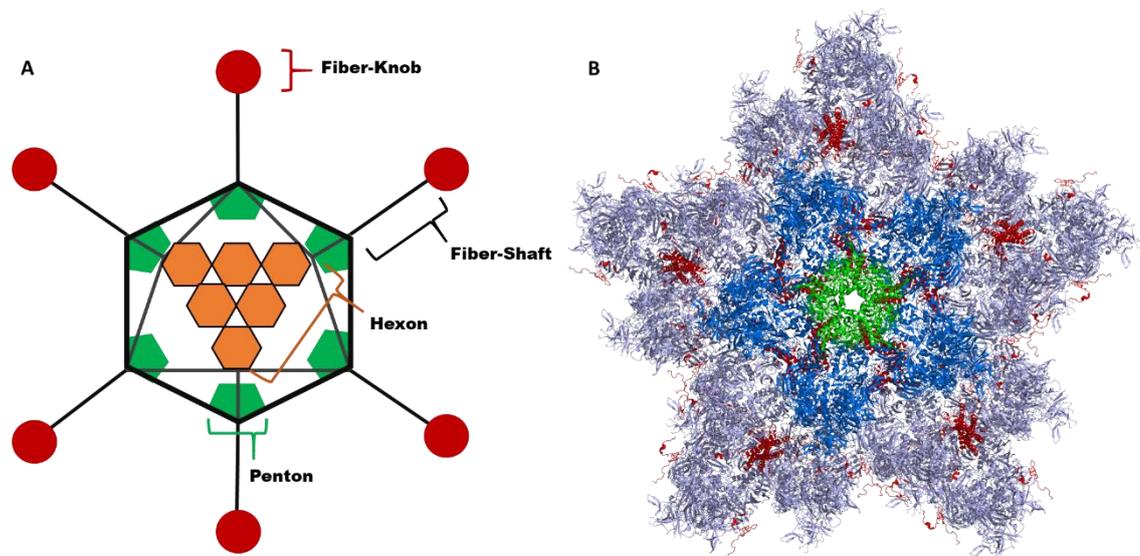


Figure 1-1 Cartoon View of Adenovirus highlighting the major capsid proteins, labelled. Structural view of an Adenovirus vertex modelled from CryoEM structure 6B1T. Penton Protein is in Green, the hexon proteins immediately adjacent to the penton are in dark blue, the remaining hexons are in light blue. Minor capsid proteins are shown in Red, Fiber is not shown. determinant of immunogenicity⁴⁷.

The penton base protein is pentameric complex and is found at the 12 vertices of the icosahedral capsid, which are the five-fold symmetry axes of the main capsid^{46,48}. This enables tessellation of the hexons into the 3-dimensional icosahedron. The surface of the penton interacts, non-covalently, with the N-

terminal domain of the trimeric fiber-protein. From the N-terminal domain the fiber extends away from the capsid to form the semi-flexible shaft domain⁴⁹. This then forms a globular domain, the fiber knob, at the C-terminus. The fiber knob protein consists of a central β -sandwich (sometimes called a β -barrel) fold consisting of approximately 10 β -strands (A-J) which are mostly conserved between species. The β -stands are joined by semi-flexible loops with highly variable sequences, named after the β -stands they are between (e.g. the HI loop falls between β -stand H and I). This fiber knob protein is the primary tropism determinant.

In the course of infection, the fiber knob protein forms the primary interaction with the cell surface. There are several possible receptors for the fiber knob including CD46 (Membrane Complement Protein, MCP)^{50,51}, Coxsackie and Adenovirus Receptor (CAR, CXADR)^{52,53}, Desmoglein 2 (DSG2)^{54,55}, and sialic acid^{27,56-59}. The identity of the receptor for the fiber knob on the cell surface is dependent upon the adenovirus species. The penton protein then forms a secondary interaction with integrins $\alpha\beta 3/5$ which triggers endocytosis^{60,61}.

The hexon is also capable of acting as an indirect tropism determinant. When the adenovirus enters the blood it is exposed to numerous proteins, including blood coagulation factors. Coagulation factor X (key to adenovirus hepatic tropism, described in section 1.4.6.) has been shown to bind to HAdV-C5 hexon with high affinity, and then bind heparan sulfate proteoglycans (HSPGs), in turn^{62,63}. This makes FX an indirect receptor for the adenoviruses capable of interacting with it.

The adenovirus capsid also includes several minor proteins. These include pIIIa, pVI, pVIII, and pIX. Proteins pIIIa and pVIII act as a kind of cement, with each capsid containing 60 and 120 copies, respectively. These two proteins are in the capsids interior and stabilise interactions between the penton base and hexons to assist capsid formation⁴⁶. Protein VI is also interior and plays a part in several processes including endosomal escape and importing hexon proteins to the nucleus for virion assembly^{46,64}. Protein IX forms a unique triskelion shape and has 240 copies on the capsid exterior which helps to stabilise the facets of the icosahedron^{46,64}.

1.4 Disrupting natural adenovirus tropisms to reduce off-target infection

Adenoviruses are not tumour selective during infection, in their natural pathogenic state. The wildtype viruses cause transient and generally non-life-threatening infections of the respiratory and gastro-intestinal tracts, as well as ocular infections. This is due to the native adenoviruses tropisms, which result in infection and spread of virus within healthy tissues. If left unrefined, adenovirus-based therapies have little uptake in cancerous cells and exhibit reduced efficacy due to sequestration of the virus in non-target tissues, clearance of the virus prior to delivery to the target tissue, and consequently result in dose limiting toxicities^{14,65}. Therefore, to achieve maximum therapeutic efficacy with minimal off-target toxicity, extensive refinement of adenovirus tropism is required to tailor the virotherapy into a cancer selective agent.

Adenoviral serotypes bind different receptors, reflecting the diversity of natural adenoviral pathogenicity as reviewed by Ghebremedhin⁶⁶. Of note is the fact that adenoviruses capable of binding the same receptors may do so with different affinities, exemplified by HAdV-B11 and HAdV-B21, the latter has substantially lower affinity for, and a different binding mode to CD46⁵¹. Receptor tropisms can be predicted from sequence alignments and structural informatics, but often remain biologically unconfirmed. The modifications required for effective viral detargeting strategies depend upon the choice of adenoviral serotype used, and a detailed mechanistic understanding natural virus: host receptor interactions is therefore essential to enable rational modification of the capsid proteins.

The information about adenovirus receptors described in this section is summarised in table 1-1.

Table 1-1 Adenovirus receptors, the residues shown to facilitate receptor interactions, and demonstrated tropism ablating mutations. *Prototype viral receptor refers to the adenoviral serotype and protein used in the receptor interaction study for which the binding residues and mutations are described. Named mutations are in bold with the mutations in brackets. amino acids are described with single letter code. Δ indicates a deletion mutation, while substitution mutations are described with the original residue letter codes preceding the residue numbers and then the respective substituted amino acid code.*

Receptor	Prototype Adenovirus Attachment Protein	Receptor Binding Residues	Previously Demonstrated Tropism Ablating Mutations	References
CAR	HadV-C5 – Fiber Protein knob	A406; S408; P409; R412; Y477; R481; L485; Y491	KO1 (SP408-409EA); KO2 (ΔVK441-44s); KO3 (R460E); KO4 (ΔGK509-510); KO5 (ΔGT538-539); KO8 (N468T); KO9 (V466H); KO10 (P505A); KO11 (Δ404-581 Whole region) ΔTAYT (ΔTAYT489-492)	Jakubczak et al 2001 ¹¹⁴ Roelvink et al 1999 ¹¹²
CD46	HAdV-B35 – Fiber Protein knob	F132; N133; T136; R169; M170; S172; N194; E192	F242; R279; S282	Wang et al 2007 ¹²⁶
DSG2	HadV-B3 – Fiber Protein Knob	N186; V189; S190; D261; F265; L292; L296; E299	N186D; V189G; S190P; D261N; F265L; L296R; E299V; ND186-261DN; ΔD261+L296R; NDL186-261-296DNR.	Wang et al 2013 ¹⁵²
GD1a/Sialic acid	HadV-D37 – Fiber Protein knob	Y308; Y312; P317; V322; K322	None reported	Nilsson et al 2011 ¹⁶⁰
Blood Coagulation Factor X	HadV-C5 – Hexon Protein HVR's	HVR regions 3; and 7 (Individual residues not clearly defined).	Ad5HVR48 (Ad5 with the HVR's of Ad48) HVR5-BAP (71aa BAP peptide insert) HVR5* (TE268-269AT); HVR7* (ITEL420-422-423-425GNSY); E451Q	Waddington et al 2008 ¹⁸⁷ Kalyuzhniy et al 2008 ¹⁸⁸ Alba et al 2009 ¹⁸⁹

HSPG	HadV-C5 – Fiber Protein shaft	KKTK91-94	S* (KKTK91-94GAGA); KKTK91-94RGDK	Paolo et al 2007, Kritz et al 2007 ^{183,184}
Integrin	HAdV-C5 – Penton Protein	R340, G341, D342	RGE (D342E) EGD (R340E)	Bai et al 1993 ²⁰¹ Henning et al 2005 ²⁰²
MARCO	HadV-C5 – Hexon Protein	HVR1; implied but not conclusively determined	None reported	Stichling et al ²⁰³

1.4.1 CAR (Coxsackie and Adenovirus Receptor)

Perhaps the best studied adenovirus tropism is Coxsackie and Adenovirus Receptor (CAR), which is widely expressed in the tight junctions of epithelial tissues^{67,68}. CAR was originally shown to be the receptor for adenoviruses HAdV-C2 and HAdV-C5⁶⁹. It has since been shown to be a receptor for a diverse range of adenoviruses across the species⁷⁰. This includes members of species A, C, D, F, and G^{70,71}. Recent work suggests that CAR is likely to be a receptor for most, if not all, of the species D adenoviruses⁷².

Crystal structures have been obtained for HAdV-A12⁵² and HAdV-D37⁷³ in complex with the CAR-D1 domain, as well for canine adenovirus type 2 (CAV2)⁷³. Oddly, the CAR bound fiber knob structure has not been described for the prototypical CAR interacting HAdV-C5. However, the HAdV-C5: CAR interaction has been characterised in detail through mutational analysis of the

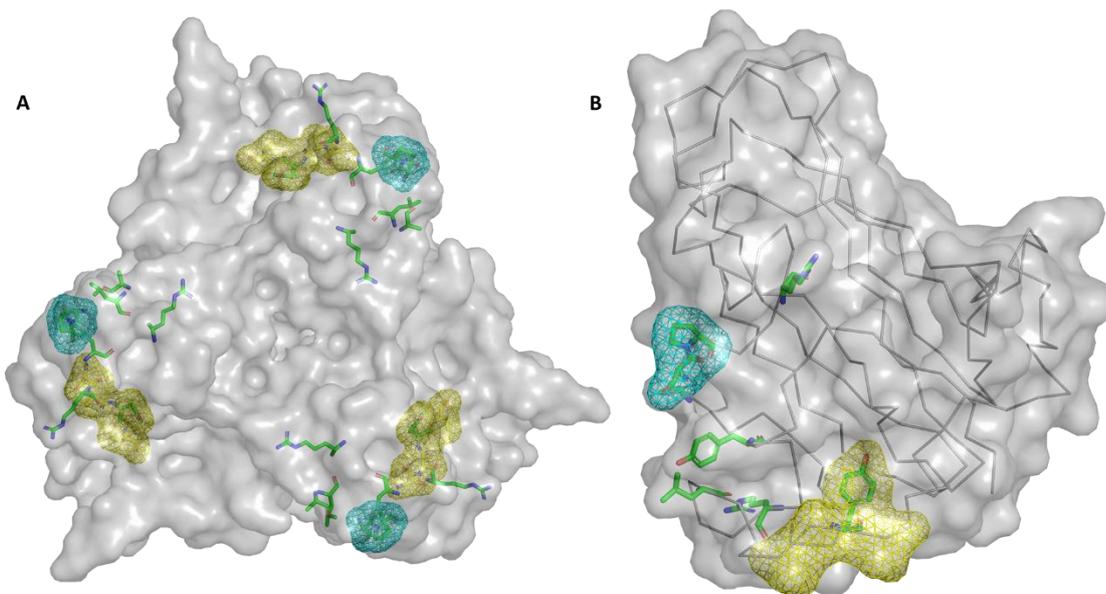


Figure 1-2 CAR interacting residues within the HAdV-C5 fiber knob domain. Known CAR interacting residues are shown as green sticks. The blue and yellow mesh shows the surface of the KO1 and Δ TAYT mutations, respectively. The trimeric structure (A) is derived from the original fiber knob monomer from PDB: 1KNB (B).

HAdV-C5 Fiber knob domain⁷⁴⁻⁷⁷.

Mutations to ablate the interaction between adenovirus and CAR have been well explored in the context of Ad5^{74,76,77}. The most efficacious, and therefore most widely used, of these are the KO1 (Fig1-2, the 1st of 10 CAR binding ablation mutations in the same study), KO2, and Δ TAYT mutations (Fig1-2)^{74,76}.

KO1 has been shown to almost completely disrupt the ability of HAdV-C5 to bind CAR both *in vivo* and *in vitro*, though it is not sufficient to alter the native viral hepatic tropism following intravascular administration^{78,79}.

The Δ TAYT mutation removes residues 489-492 in the FG loop of HAdV-C5 (Fig1-2). The TAYT motif was predicted to be important to receptor binding due to its high conservation between 14 adenovirus species (Ad2, 4, 5, 8, 9, 12, 15, 17, 19, 28, 31, 37, 40 long, and 41 long)⁷⁴, suggesting that this mutation may be effective in other Adenoviruses. The $\alpha\beta 6$ targeted Ad5-3 Δ -A20T oncolytic virus shows the efficacy of this mutation in CAR binding ablation *in vivo*, though as with the KO1 mutation, ablation of CAR binding alone is insufficient to eliminate accumulation in the liver^{80,81}. However, CAR-binding ablation may combat sequestration of adenovirus in the blood by preventing binding to CAR expressing erythrocytes, with the KO1 mutation being shown to prevent hemagglutination as evidence to support this further^{60,82,83}.

1.4.2 CD46/MCP (Membrane Cofactor Protein)

Cluster of Differentiation 46 (CD46), also known as Membrane Cofactor Protein (MCP), is a complement regulatory protein expressed on the surface of all known nucleated cells⁸⁴. This nearly ubiquitous protein has been shown to be important to the pathogenicity of several pathogens^{85,86}. Its native function is to protect against complement activity, and it has been implicated in the induction of regulatory T-cells. This is all covered in the review by Yamamoto *et al*⁸⁷.

Presently the only full length crystal structure of the CD46 extracellular domain has been determined in complex with the fiber knob domain of Adenovirus serotype 11⁸⁸. The SCR1 and SCR2 domains of CD46 have been crystallised in complex with HAdV-B21⁵¹. This is one of several Species B adenoviruses which have been shown to utilise CD46 as their primary cell entry receptor^{51,89,90}. Crystallographic structures are available for Ad11 and Ad21 in complex with CD46. The authors of the Ad21 complex structure suggest that there may be two modalities of CD46 interaction governed by the length of the HI and DG loops^{51,88,90}.

Mirroring the situation with the Ad5:CAR interaction, there is no ligand bound structure of the quintessential CD46 interacting fiber knob, HAdV-B35. Despite

this, the interaction is well explored, and the individual key binding residues characterised (Fig1-3)⁹¹. The analysis of CD46 contact residues reveals the locations of critical mutations which have been used to abrogate CD46 interaction (Fig1-3)⁹¹.

At the time of writing there is no published development of therapeutic vectors for which CD46 binding has been ablated. This is curious when compared to the widespread adoption of de-targeted CAR binding adenoviruses. This may be a result of the perceived benefit of a universal entry pathway and regulation of replication post infection. While this approach has the benefit of a simpler vector design strategy, it has significant drawbacks. These include CD46 usage by adenoviruses being associated with reduced expression of pro-inflammatory cytokines and immunosuppression, as well as dose limiting toxicities resulting from larger adenovirus titres required to overcome off-target infection and reach the tumour present difficulties⁹²⁻⁹⁴. Immunosuppression is of particular concern with the knowledge that immune stimulation and the breaking of tumour immune tolerance are key mechanisms by which oncolytic adenoviruses are effective at eliminating the tumour^{2,95}.

While CD46 binding ablating mutation have not been identified, the differences observed between serotypes offer clear mechanisms by which CD46 interaction may be inhibited. The involvement of four loops has been observed for HAdV-B7, B11, B35, and B21 in binding CD46 domain SCR1 and SCR2^{51,89,90}. HAdV-B21 was shown to cause CD46 to bend further to accommodate its longer DG-loop. This was suggested as the reason for its lower CD46 binding affinity ($K_D = 284\text{nM}$) compared to HAdV-B11 ($K_D = 13\text{nM}$) and HAdV-B35 ($K_D = 19\text{nM}$)⁵¹. Therefore, it is logical that longer DG-loops would inhibit CD46 binding further, as they would cause it to bend to a point where it was unable to contact the interacting loops.

CD46 utilising adenoviruses remain a popular choice for the development of oncolytic virotherapies since CD46 is upregulated in numerous cancers, though CD46 is far from unique to cancer as it is present on the surface of all nucleated cells⁹⁶. An example is ColoAd1/Enadenotucirev (EnAd), an evolved and selected for Ad11p/Ad3 chimera, generated by recombination and selection of adenoviruses rather than rational modification, and has progressed to phase I/II

clinical trials⁹⁷. EnAd possesses promising oncolytic properties derived from several mutations, including E3 and E4 region deletions^{97,98}, and is purported to

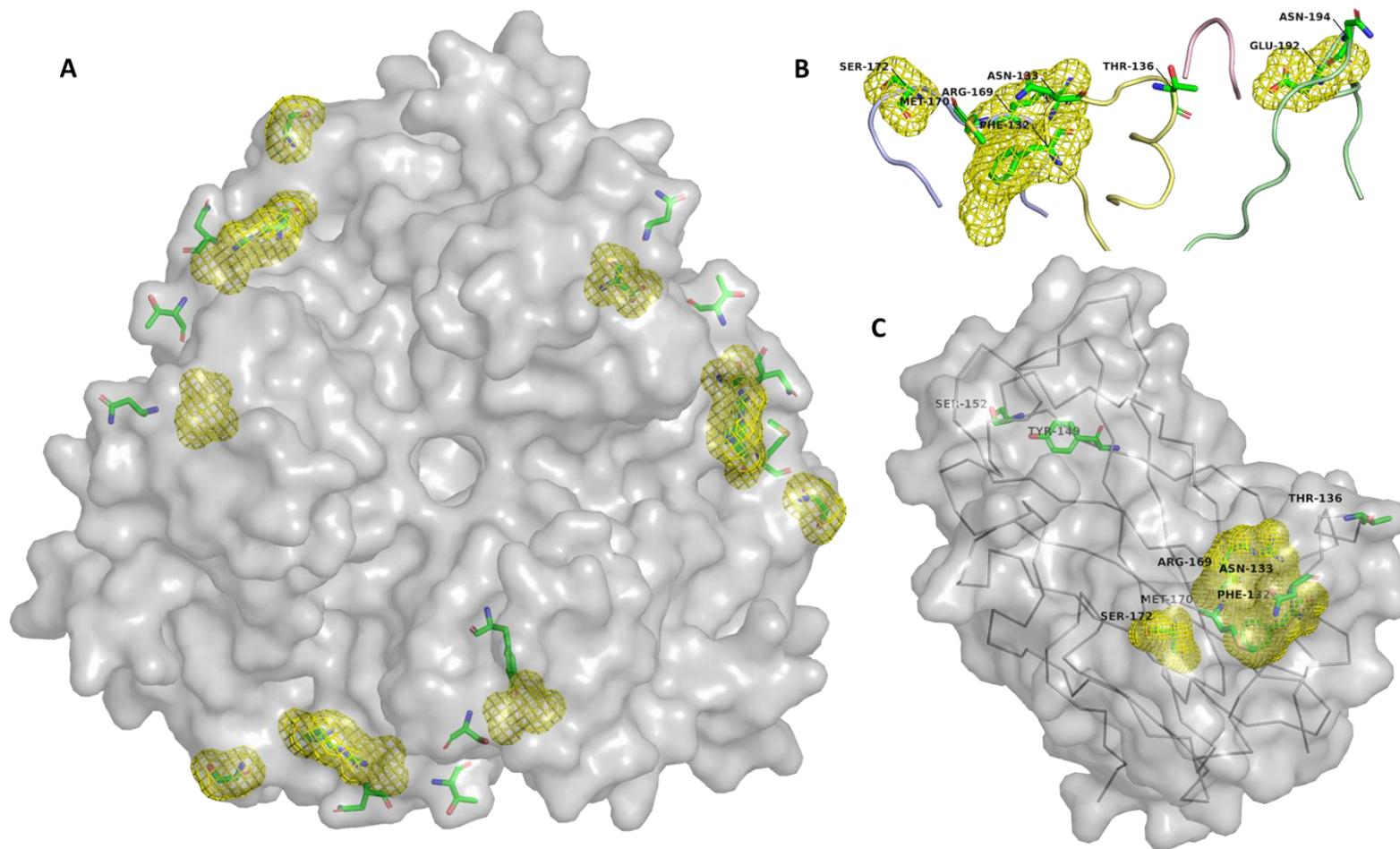


Figure 1-3 CD46 interacting residues, and known mutation sites, within the HAAdV-B35 fiber knob domain: *The key residues which interact with CD46 are shown as green sticks. The yellow surface shows the region in which known mutations which abrogate CD46 interaction occur. The trimeric HAAdV-B35 fiber repeats this interface 3 times (A). A detailed view of the 4 loops which interact with CD46, HI (Blue), DG (Yellow), GH (Red), and IJ (Green) shows how the interface is formed in space (B) and the monomeric HAAdV-B35K shows where this is located on the monomer (C). Structure from PDB: 2QLK.*

have a primary tropism to both CD46 and Desmoglein 2^{99,100}. Though the affinities have not been fully defined, the fiber protein drives the affinity of HAdV-B11 for CD46 and DSG2^{55,91,101}, is 100% identical to that of EnAd at an amino acid sequence level¹⁰². It therefore seems reasonable to assume that EnAd will behave similarly in terms of tissue tropism.

It has previously been suggested that the species D adenoviruses HAdV-D26 and HAdV-D48 may also utilise CD46; vectors derived from these viruses are currently under development as HIV and Ebola vaccines^{41,103–105}. This alleged tropism has been challenged, with other papers suggesting CAR as the primary receptor^{106,107}. However, recent work shows that the primary tropism determining fiber knob proteins of HAdV-D26 and HAdV-D48 are incapable of binding to CD46, while they have only a weak affinity for CAR⁷². The true receptor for these fiber knob proteins appears to be sialic acid⁵⁸.

1.4.3 Desmoglein 2

Desmoglein 2 (DSG2) is part of the Desmoglein subfamily of Cadherin junctional adhesion proteins and is one of four desmoglein isoforms. DSG2 acts as a cell entry receptor for species B Adenoviruses 3, 7, 11, and 14⁵⁵. DSG2 requires knob domains to form a constellation to mediate cell entry as shown by studies of whole virus, dimerised knob domains, and penton dodecahedrons (a natural particle produced by HAdV-B3 composed of 12 penton-fiber complexes which is capable of opening DSG2 mediated cell junctions)^{108–112}.

Similar to CD46 utilising viruses, efforts have focused on exploitation of this tropism rather than abolition of DSG2 binding. Work from the lab of David Curiel generated an HAdV-C5 pseudotyped with the fiber knob of HAdV-B3 (HAdV-C5/B3K) which was shown to efficiently infect prostate cancer cells¹¹³. Development of the HAdV-C5/B3K virus led to the CGTG-102 virus, engineered to express granulocyte-macrophage colony stimulating factor (GM-CSF)¹¹⁴. CGTG-102, now branded as ONCOS-102, has progressed into phase 2 clinical trials in colorectal cancer and pleural mesothelioma having shown safety in a small phase I trial, while a similar virus showed efficacy in a phase I trial of ovarian cancer^{114–118}. ONCOS-102 has recently announced the recruitment of 31 patients to a phase Ib/II clinical trial of unresectable pleural mesothelioma¹¹⁹.

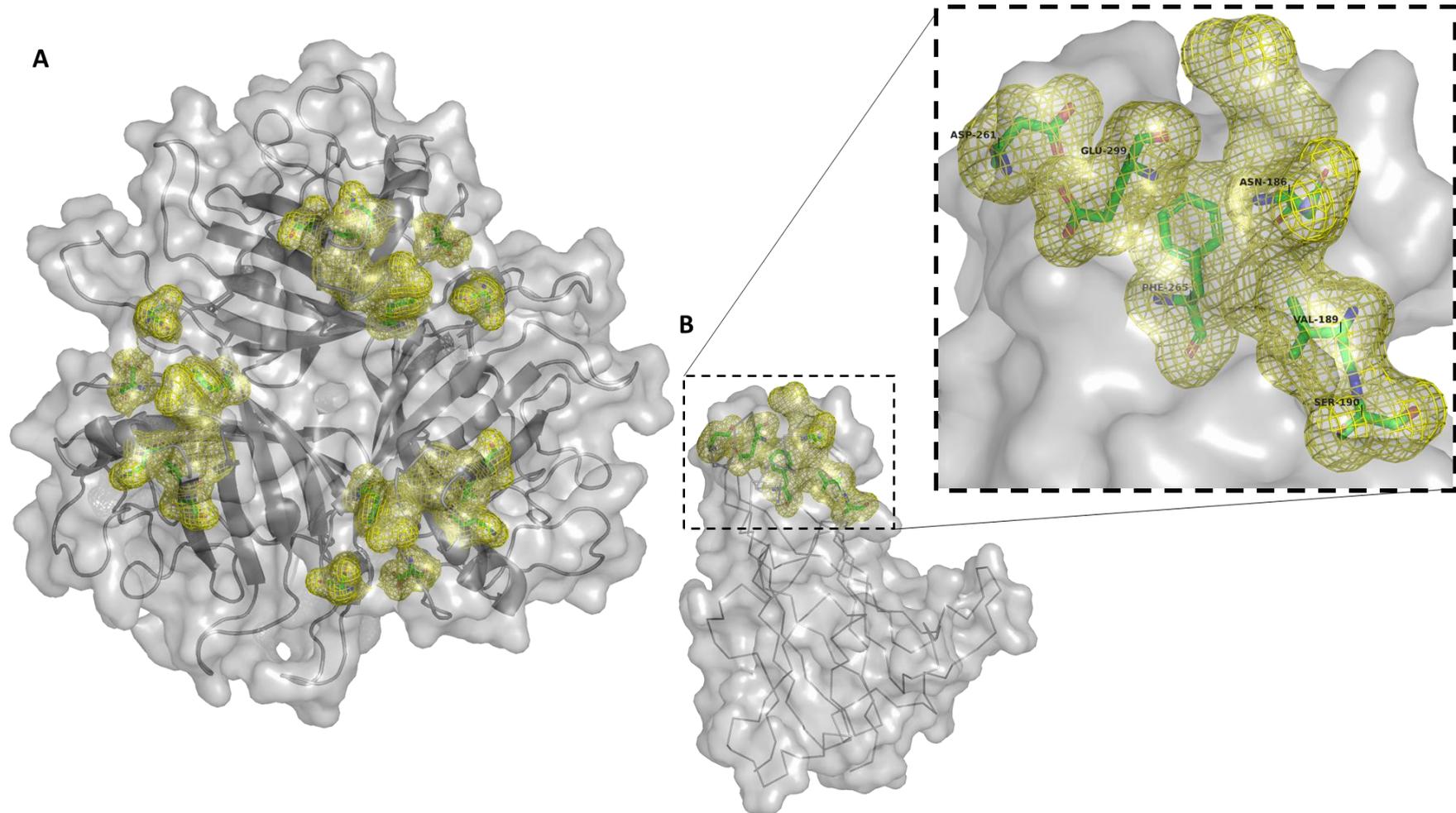


Figure 1-4 DSG2 interacting residues, and known mutation sites, within the HAdV-B3 fiber knob domain. *The key residues which interact with DSG2 are shown as green sticks on the trimeric fiber knob (A) and on the monomer (B). The yellow surface shows the region in which known mutations which abrogate DSG2 interaction occur. Structure from PDB: 1H7Z*

Work has been performed to increase the affinity of Ad3 fiber knob for DSG2 in pursuit of a therapeutic molecule capable of loosening epithelial cell junctions to improve tumour penetration, culminating in the JO-4 molecule^{110,120-122}. In the process of finding these affinity enhancing mutations, the lab of André Lieber identified several mutations capable of reducing or ablating the Ad3-DSG2 interaction (Fig1-4). When mapped to the Ad3 fiber knob crystal structure all of these mutations are at the extreme distal end suggesting the location of the binding interface which may overlap with that of GD1a glycan (see section 1.4.4)^{123,124}. This has now been confirmed with the recent publication of the cryo-electron microscopy (CryoEM) structure of HAdV-B3K in complex with DSG2¹²⁵. This shows a loop invading the apical indentation of HAdV-B3 where the sialic acid would otherwise bind, and confirms a previous publication asserting that, unlike CAR and CD46, DSG2 has a 1:1 stoichiometry with HAdV-B3K, though a minor population of 2:1 stoichiometry was also shown⁵⁴.

It has been suggested that Adenovirus binding DSG2 can induce an Epithelial to Mesenchymal Transition (EMT) phenotype, and induce transient opening of cell junctions^{55,126}. Presumably this facilitates viral spread through the tissue. However, EMT has also been linked to cancer metastasis, though it is unclear if EMT is strictly necessary for metastasis and remains an area of investigation¹²⁷⁻¹³⁰. DSG2 interacting molecules may be capable of inducing EMT/MET (Mesenchymal Epithelial Transition) and treatments based on the ability to bind DSG2 must therefore be monitored closely for this potential limitation, especially in the context of a cancer targeted oncolytic virus.

1.4.4 Sialic acid, Glycan interaction, and GD1a

Sialylated glycans are expressed on the surface of innumerable proteins, and are the receptor for many viruses¹³¹. Sialic acid makes an attractive target for viruses as they are conserved, presented on the cell surface, and available in large numbers. Several adenoviruses have been shown to utilise sialic acid as a cell entry receptor. These include HAdV-G52 which uses polysialic acid as a receptor to the lateral region of its short fiber knob (HAdV-G52SFK)⁵⁷, as well as Turkey Adenovirus 3 (TAdV-3)¹³² and CAV2⁷³ which also attach to sialic acid in the lateral regions of their fiber knob proteins. However, the most widely

observed mode of sialic acid binding is in the central apical depression of the fiber knob protein seen in many species D adenoviruses^{27,58}.

1.4.5 Polysialic acid

A recent study from the Arnberg lab reveals HAdV-G52, which has two different fiber proteins (SFk – Short Fiber, LFk – Long Fiber), bind to polySia (α -2,8-linked poly sialic acid) via its SFk, the LFk binds to CAR^{57,71}. The authors demonstrate a primary hydrogen-bond interface with the first sialic acid of the polySialic acid glycan chain and use molecular dynamics simulations to argue, convincingly, that transient electrostatic contacts stabilise the interface, explaining the increased affinity with polySia chains ≥ 3 units.

PolySia is a rare post-translational modification (PTM). Expression is largely restricted to neurological tissues of the hippocampus, olfactory bulb, and hypothalamus in healthy adult brains^{133–135}. It has been proposed that polysialylated protein is involved in the development of other organs via polysialylated NCAM (Neuronal Cell Adhesion Molecule), and also in innate immunity via polysialylated CCR7. This are just two of the reported polySia carrier proteins^{136,137}. It is not known if HAdV-G52 infection is limited to particular polySia carrier proteins. While the only known isolation of HAdV-G52 is from a patient with gastroenteritis the expression pattern of polySia suggest a potential neurological tropism¹³⁸. However, much is unknown about polySia and the HAdV-G52 adenovirus, making predicting its behaviour difficult at best.

Given polySia is expressed on NCAM, a neurological tropism is of potential concern when engineering a virus for cancer virotherapy in the context of off-target infection. However, several cancers including glioma, astrocytoma, neuroblastoma, and NSCLC (Non-Small Cell Lung Cancer) over-express polySia making it a potentially interesting candidate therapy for these cancers^{139–145}. While general polySia expression within the brain seems likely to generate off target effects this could be controlled post viral entry using strategies such as tumour specific promoters to control replication. This tropism could have potential for usage in vectors for NSCLC or neuroendocrine tumours when used in conjunction with the appropriate selectivity mutations.

The paper by Lenman *et al*, describes a receptor binding modality reliant upon transitory interactions. When the primary sialic acid contacts the first sialic acid it forms a weak interface, which is then stabilised by roving electrostatic interactions. This could make engineering of this tropism extremely challenging as any alteration to ablate the interaction would have to alter the electrostatics of the entire region. Conversely, any effort to enhance the interaction would have to preserve the overall electrostatic character of the region while finding mechanisms to enhance polar interactions with the primary sialic acid residue⁵⁷.

1.4.6 Sialic acid binding in the species D adenoviruses

The most widely observed and best studied mode of adenovirus binding to sialic acid is typified by HAdV-D37 which binds to sialic acid in the apical depression of the fiber knob²⁷. HAdV-D37K, which has an identical fiber knob protein to HAdV-D64 (formerly HAdV-D19a), HAdV-D19pK, and HAdV-D26K have all been crystallised in complex with sialic acid^{27,58}.

The interaction of HAdV-D37K with sialic acid is well established^{27,56,146,147}. HAdV-D37K has high sequence identity to several other fiber knob proteins, all of which are belong to viruses known to cause EKC. Consequently, it has been assumed that the similar EKC causing viruses, HAdV-D8, HAdV-D53, HAdV-D54, HAdV-D56, and HAdV-D64, use sialic acid for infection. This has recently been confirmed for all these viruses with the exception of HAdV-D56 which was shown to infect human conjunctival epithelial (HCE) cells even in the absence of sialic acid on the cell surface⁵⁹.

The recent discovery that HAdV-D26K is capable of binding to sialic acid and using it to initiate a productive infection changes our understanding of sialic acid usage in the species D adenoviruses⁵⁸. HAdV-D26K is situated outside of the clade occupied by the EKC causing adenoviruses⁷². This suggests that sialic acid usage in the species D adenoviruses may not be as limited as previously thought and may extend across the species D adenoviruses.

There are an array of sialic acid terminating glycans which could act as receptors for adenovirus infection. HAdV-D37 has been shown to prefer the disialylated GD1a glycan for cell entry (Fig1-5)¹⁴⁷. The asymmetric interaction

with the fiber knob means only one copy of the GD1a can interact per fiber knob domain, in

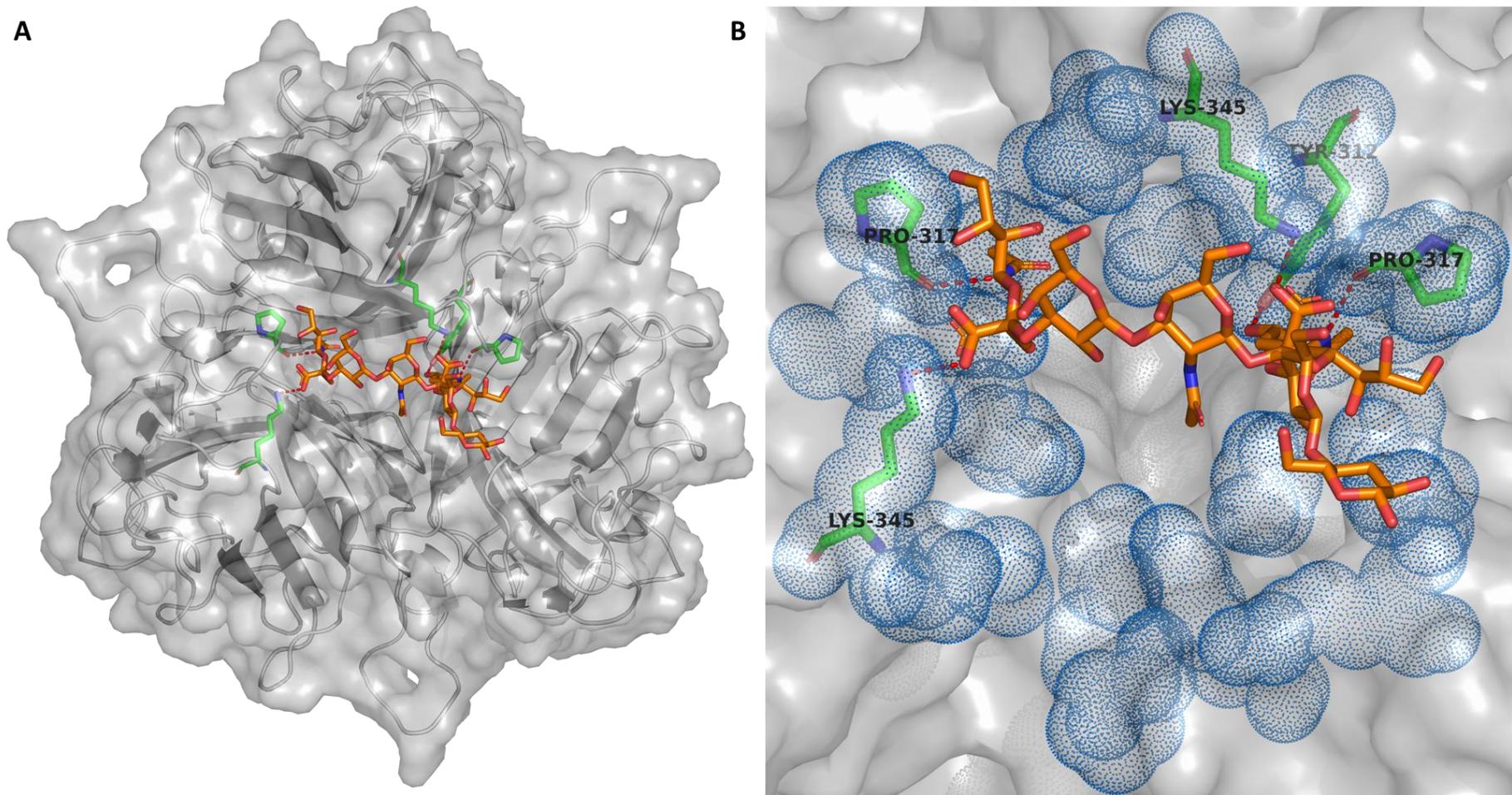


Figure 1-5 GD1a/Sialic Acid interacting residues within the HAdV-D37 fiber knob domain. Key residues forming the GD1a-Ad37Fkn interaction are shown as green sticks, with the GD1a in orange, on the HAdV-D37 fiber knob trimer (A). Hydrogen bonds are shown by red dashes. While the interface can occur in three orientations, only one set of interacting residues is shown. On the close-up view of the fiber knob apex (B) the blue dots the surface of all residues shown to be able to interact with GD1a or support the interaction, seen to create a large apical binding pocket. Structure from PDB: 3N0I.

contrast with the three copies of CAR and CD46. In addition to the conservation of key binding residues, preincubation of HAdV-D8, D9, D19p, or D37/64 fiber knob protein with soluble GD1a inhibited virus binding to human corneal cells to varying degrees¹⁴⁷. However, it seems unlikely that GD1a glycan is the only glycan which these viruses can use as an entry receptor, especially since HAdV-D26 is more strongly associated with respiratory infection. Interestingly HAdV-G52, HAdV-D37, and CAV2 have both been modelled in complex with CAR and sialic acid simultaneously, demonstrating the ability of adenoviruses to bind two different classes of receptor at once^{71,82}.

Most research regarding sialic acid binding adenoviruses has been in pursuit of an effective antiviral treatment for EKC. As a result efforts have focused on the design of sialic acid mimetic inhibitors, similar to those used to treat influenza^{148,149}. Since genetic variants are of little help in treating naturally acquired infections, no sialic acid/GD1a binding ablation mutations have yet been reported in adenoviruses. Nevertheless, Tyr312 and Lys345 residues are shown to be critical for GD1a binding in HAdV-D37 and are completely conserved in other species D adenoviruses^{72,147}. These and the supporting Tyr308, Pro317, and Val322 residues present obvious targets for mutation, but remain untested (Fig1-5).

1.4.7 Coagulation Factor X (FX), Heparan sulfate proteoglycans (HSPGs), and IgM

HSPGs are expressed on all cell surfaces and represent the receptor for myriad viruses including those with oncogenic properties (except Epstein-Barr virus)^{150,151}. Adenoviruses have been shown to interact with HSPG's through differential mechanisms. HAdV-C2, B3, and C5 have all been shown to interact with HSPG's in a sulfation dependent manner (higher sulfation increases binding). The interaction is mediated by a BBXB motif (B is a basic residue, X is a hydrophobic residue) typified by the KKTK motif which occurs in the HAdV-C5 and C2 fiber-shafts. BBXB is one of two consensus sequences proposed for Heparan sulphate glycosaminoglycans (HSGAG) interaction, the other being BBBXB¹⁵²⁻¹⁵⁵.

The idea that the KKTK motif and subsequent HSPG interaction via the fiber-shaft was solely responsible for trafficking of adenovirus to the liver has been dispelled by studies utilising viruses with mutated or deleted KKTK motifs. Chimerism between the fiber proteins of HAdV-A31 and F41 (CAR interacting Adenoviruses lacking the KKTK motif), the S* modification (HAdV-C5 residues $_{91}\text{KKTK}_{94} \rightarrow _{91}\text{GAGA}_{94}$), and the $_{91}\text{KKTK}_{94} \rightarrow _{91}\text{RGDK}_{94}$ mutation to substitute KKTK for an additional integrin interacting motif, all abolish HSPG binding, but hepatic tropism is retained^{156–158}. These studies also demonstrate the fiber-shaft KKTK motif is needed to create enough flexibility to be able to bend and facilitate engagement between the viral penton protein and cellular $\alpha\beta 3/5$ integrins. Mutation of these key residues (as in the S* mutation) render the virus fiber-shaft inflexible compromising the virus's fitness.

It has now been shown that adenovirus liver transduction is mediated by a high affinity, Ca^{2+} ion dependent, interaction between the major capsid protein, hexon, and blood coagulation factor X (FX)⁶². CryoEM was able to identify hexon hypervariable regions (HVR) 3, 5, and 7 as critical interaction determinants^{159–161}. Mutations in HVR5 and 7 have been shown to efficiently abrogate HAdV-C5 trafficking to the liver by ablating the hexon: FX interaction. This has been achieved by both rational modification of the hexon, and by pseudotyping of the HAdV-C5 HVRs with the HVRs of non-FX interacting adenoviruses, namely Ad26 and Ad48^{63,160,162,163}. The HVR7 mutant was especially effective in detargeting HAdV-C5 from the liver *in vivo*¹⁶⁰.

FX may also shield HAdV-C5 from complement mediated attack via inhibition of IgM binding¹⁶⁴. Therefore, it is plausible that ablation of this interaction may result in an increase rate of clearance of the non-FX binding virus from the body, potentially reducing therapeutic efficacy¹⁶⁵. However, mutation of hexon protein to ablate FX interactions could also be made to inhibit IgM binding. Albeit, this may preclude the potential beneficial effects provided by steric inhibition of hexon interaction by large molecule binding (such as FX), which may represent a broadly effective mechanism for inhibiting immune recognition. Such a mechanism was described by Schmid *et al* (discussed later in this review)¹⁶⁶.

Not all adenoviruses bind FX, proving that interaction with FX is not a requirement for immune evasion, and this immune evasion could be achieved through incorporation of a hexon or hexon hypervariable region(s) derived adenoviruses that are unrecognized by the immune system¹⁶⁷. Therefore, whilst ablation of FX binding may have effects on adenovirus clearance the virus must be considered individually in the context of its native FX binding ability and seroprevalence in the population. Consequently, all engineered viruses must be considered as unique entities and tested for these properties since the interplay of HVR regions with clearance, immune protection, and HSPG binding is difficult to predict.

Hepatocyte mediated viral uptake occurs via an indirect interaction. HAdV-C5 hexon binds to FX, in turn the FX binds to cellular HSPGs. The tripartite complex is then internalised by the hepatocyte¹⁶⁸. This pathway was found to be particularly dependent on O-linked HSPG sulfation state, which are abundant on hepatocytes, with higher sulfate levels being linked to more efficient uptake¹⁶⁹.

Despite the strong interaction between FX and HSPGs the fiber protein continues to influence the final tissue tropism. HAdV-C5/B35K pseudotyped viruses exhibited differential liver tropism to HAdV-C5 when FX interaction was ablated¹⁷⁰. The mechanism underlying this effect remains undetermined, but pH dependent effects on endosomal transport may relate to differential properties of the HAdV-B35K vs HAdV-C5K in endosomal escape, rather than a direct interaction with HSPG. Alternatively the relative receptor abundance of CD46 or CAR may overcome part of the alternative tropism and alter the tropism.

Most vector development has focused on ablation of HSPG binding, but the HSPG interaction has been exploited in therapeutic development. The HAdV-C5.pk7-Δ24 virus contains a poly-lysine motif in the C-terminus of the fiber to facilitate CAR independent infection via HSPG¹⁷¹.

These studies highlight that FX interaction must be eliminated to effectively disrupt HSPG interaction. However, the effectiveness of the poly-lysine modification, and the ability of peptides to form electrostatic interactions implies that absolute abolition of HSPG binding may be unfeasible given the adenovirus capsid has an inherent charge^{171,172}. The electrostatic mode of HSPG binding

means any patch of positive charge on the capsid represents a potential binding site. Given the variability of adenoviruses net electrostatic potential it seems likely that the robustness of the aforementioned mutations will vary depending upon serotype¹⁷³. However, if interactions are brought to sufficiently low affinity this is unlikely to have *in vivo* relevance to the tissue tropism.

1.4.8 Integrins $\alpha v\beta 3$ and $\alpha v\beta 5$

Integrins $\alpha v\beta 3$ and $\alpha v\beta 5$ were shown to be required for internalisation of HAdV-C2 in 1993 via interaction with the penton base, which initiates clathrin dependent endocytosis^{174,175}. Endocytosis requires PI3K (Phosphatidylinositol-3-OH) Rho GTPase reorganisation of the actin cytoskeleton^{176,177}. Integrins interact via a canonical RGD motif which is widely conserved despite occurring within a hypervariable region of the penton base protein⁶¹. Mutation of the RGD motif to RGE is enough to prevent efficient infection; RGE mutant HAdV-C2 virus is shown to accumulate on the cell surface without internalising. This mutation has been shown to be effective in an HAdV-C5 background^{61,77,174,178}. Alternatively, there is the EGD mutation which achieves similar ends¹⁷⁹. This integrin mediated endocytosis mechanism has been shown to apply across almost the whole adenoviral species, with only the Ad40 species F serotype reported to have a naturally occurring RGA rather than an RGD motif^{61,180}.

There are reports that $\alpha v\beta 1$, $\alpha M\beta 2$, $\alpha L\beta 2$, and $\alpha_{IIB}\beta 2$ may also enable adenovirus internalisation in the absence of $\alpha v\beta 3$ and $\alpha v\beta 5$, though these interactions are not as well described¹⁸¹⁻¹⁸⁵. The penton base RGD loop interacts with an NPXY motif (X is any amino acid), conserved at the C-terminus of integrin β -subunits; this motif is sometimes extended to include an N-terminal phenylalanine (FNPXY)^{186,187}. Alteration of this motif has been shown to limit adenovirus infection¹⁸⁸. Given the diversity of α -subunits in the integrin dimers which adenovirus has been observed to interact with and the conservation of the C-terminal NPXY motif in the β -subunits, it seems plausible that adenoviruses only requirement for penton base: integrin interaction is the presence of a β -subunit.

The complex between HAdV-C2 and HAdV-A12 with $\alpha\beta 5$ has been solved by CryoEM¹⁸⁹. Integrins bind the penton base at 5 RGD sites and form a continuous ring structure with $\sim 60\text{\AA}$ spacing between the 5 RGD motifs (similar to that of FMDV – Foot and Mouth Disease Virus)^{190,191}. The authors suggest that this clustering of integrins stimulates endocytosis. An updated CryoEM structure of HAdV-A12 in complex with $\alpha\beta 5$ reveals that steric hindrance limits integrin engagement to four copies per penton base, despite pentavalency¹⁸⁵. Comparison of the HAdV-C2 and HAdV-A12 structures in complex with integrin reveals the more stable conformation of RGD is in HAdV-A12, constrained by lower flexibility of the loop on which RGD resides. This enables more stable integrin interaction, presenting an opportunity to engineer tighter integrin interaction and potentially modulate the viral dependency upon fiber mediated interactions¹⁸⁹.

Whilst normally a secondary co-receptor, $\alpha\beta 5$ binding alone has been reported to facilitate efficient HAdV-C5 infection alone, and that the virus can interact with the penton base with picomolar affinity (K_D of 1.4×10^{-10} M) as determined by HAdV-C5 virus binding to CAR^{low} MDA-MB-435 cells¹⁹². It is worth noting that this study does not preclude non-integrin non-CAR means of infection and would benefit from being reassessed with surface plasmon resonance experiments to determine virus-integrin affinities in isolation. It was previously predicted that engagement of $\alpha\beta 5$ was not required for viral attachment, and while this may be true when cells bear a high affinity primary receptor the aforementioned study suggests it is not an absolute requirement¹⁷⁴.

Previously studies suggest that insufficient flexibility of the fiber-shaft would limit integrin binding as the shaft would hold the capsid away from the integrin⁴⁹. Recent CryoEM of a mutant HAdV-C5 with a short fiber protein from Ad35 (Ad5F35) indicates that this prediction was correct and short fibers do have limited flexibility. However, it does not appear to preclude binding between the penton base and integrin, and shows simultaneous interaction between short-fibered adenovirus with their primary receptor and integrins is possible¹⁹³. A recent paper on HAdV-D26 shows that the native HAdV-D26 fiber is short enough that an integrin could, theoretically, span the distance to the RGD without requiring flexibility¹⁹⁴. Thus, when engineering adenovirus it is

insufficient to rely upon the primary tropism to retarget the vector, and the penton RGD motif must be ablated to prevent off target infection. This has been shown by reduced spleen uptake in double mutated HAdV-C5 viruses, reducing splenic tropism five fold in mice only once the RGD-RGE mutation was added to HVR7 mutated virus¹⁷⁸.

1.4.9 MARCO/SR-A6 Scavenger Receptor (Macrophage Receptor with Collagenous Structure)

In vivo studies show that macrophage depleted mice have greatly reduced inflammatory response to adenovirus infection^{195,196}. This effect was shown to be due to reduced adenovirus infected MARCO⁺ macrophages local to the splenic marginal region¹⁹⁷. It is now clear that MARCO has a direct role in Adenovirus recognition and is needed for anti-adenovirus cytokine response, via the cGAS/STING pathway¹⁹⁸. This effect has now been observed in species B (Ad35), C (Ad5), and D (Ad26) Adenoviruses, and deletion experiments in the hexon imply that the site of interaction involves HVR-1 (Hyper Variable Region 1)¹⁹⁹. However, the exact nature of the MARCO: Hexon interaction remains uncharacterised. With cGAS/STING being identified as a key sensor of adenovirus infection, and innate antiviral/DNA sensing pathways becoming increasingly important to the field of oncolytic virology in general, further studies of this interaction are paramount²⁰⁰⁻²⁰³.

1.5 Retargeting adenovirus vectors using engineered receptor proteins

Once naturally occurring tropisms have been ablated it is necessary to provide an alternative means of infecting the cell; one which is specific to the tissue of interest which in this case is cancer. The choice of new cell entry receptor is severely limited by several factors. It must be expressed on the cell surface, selectively expressed/upregulated in the target cancer cells, able to facilitate adenovirus endocytosis following receptor binding, and capable of being targeted by genetic modification of adenoviral capsid proteins.

The pool of known cancer specific ligands is small, and the aforementioned restrictions make it even more difficult to find one which is appropriate. Accordingly, many studies have attempted to engineer a universal vector

platform using established ligand specific molecules, detailed in this section. Others have sought to use naturally occurring tropisms from other sources to generate chimeric viruses.

Both methodologies pose challenges. An important benefit of the chimeric retargeting approach is highlighted in a study of “knobless” adenovirus. The fiber protein was engineered to have 7 shaft repeats, a trimerization motif, and the anti-Taq polymerase Z_{taq} affibody; this generated the Ad5/R7-Z_{taq}-Z_{taq} virus based on Ad5²⁰⁴. This proof of concept study demonstrated a decrease in neutralisation by sera containing neutralising antibodies (nAbs) against adenovirus. This is not surprising considering one of the major viral capsid proteins, and a significant immunogen, is deleted. However, the ability of these viruses to evade neutralisation in some sera highlights the important observation that some patients appear to neutralise adenovirus solely via the fiber knob protein, in contrast with hexon neutralisation as is oft cited^{205,206}.

1.5.1 Single chain antibodies (scFv)

By far the most common targetable proteins are antibodies, so it is logical that many have proposed the use of scFv incorporated into the viral capsid. However, most scFv require the formation of disulphide bonds, formed in the endoplasmic reticulum²⁰⁷. Adenovirus proteins fold in the reducing environment of the cytoplasm, followed by packaging in the nucleus, creating a fundamental biosynthetic incompatibility limiting the effectiveness of this approach²⁰⁸.

One example of a successful scFv fusion protein is the Ad5FFscFv47-CMV-GFP virus. An scFv fragment of an anti-IL13R α 2 (a selective marker of glioma) mAb was fused in place of the fiber knob with a T4-fibritin trimerization motif between the scFv and shaft²⁰⁹. The resultant virus selectively infected glioma cells based on IL13R α 2 expression, but the production process demonstrates the difficulty of obtaining such a stable scFv. Selection of a viable scFv from the parental hybridoma required extensive biopanning. An attempt to solve this was by retargeting a pIX-scFv fusion protein with an ER trafficking signal²¹⁰. Whilst this solved the folding issue, the pIX-scFv fusion integrated into the virus inefficiently. The authors addressed this problem by using a single domain antibody (sdAb). Though efficacious, this also requires extensive re-engineering

of the targeting ligand, nullifying the primary advantage of scFv use: the wide availability of ligand specific reagents.

The proposed solution is scFv forming a tertiary complex, altering the scFv sequence to contain a virus specific domain as well as its ligand specificity. This is the so called “adaptor” method. An example is the addition of leucine zippers, where the scFv (still expressed from the viral genome) contains one half of the

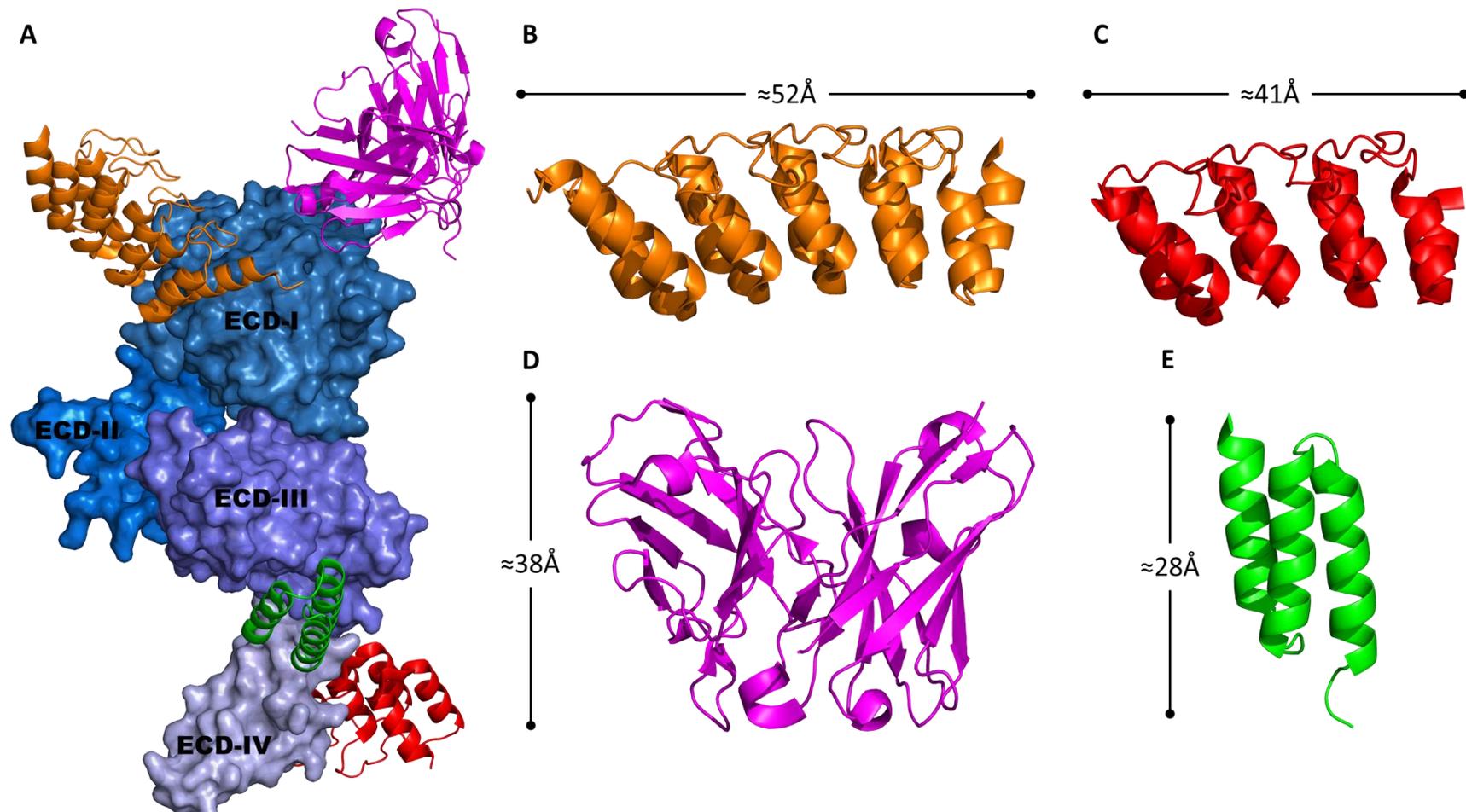


Figure 1-6 Overview of antibody-like proteins specific to HER2 which have been genetically integrated into Adenovirus. HER2 is composed of 4 domains, ECD (ExtraCellular Domain) I-IV which can be bound by different proteins (A). The DARPIN's 9.29 (B) and G3 (C) are seen in orange and red complexing ECD-I and ECD-IV, respectively. ScFv chA21 (D) is seen in purple complexing ECD-1 (this particular ScFv has not previously been integrated to Adenovirus), and affibody ZHER2 (E) is seen in green binding at the ECD-III/IV interface. All molecules are shown to scale, structures from PDB: 4HRL, 4HRN, 3H3B, and 3MZW.

pair with the other expressed in place of the fiber knob, C-terminal of a trimerization motif to retain stability²¹¹. A similar approach used a bispecific diabody, a genetic fusion of two scFv with different specificities. This molecule has specificity for CD105 (Endoglin), to target vascular tissues, and adenovirus fiber knob²¹². Another used a fusion of Ad5's native CAR receptor with an anti-HER2 scFv²¹³. A similar scFv (chA21) is seen in figure 1-6A,D, though this clone has not been integrated into Adenovirus. However, at time of writing, no adenovirus targeted by using an adaptor molecule has translated to the clinic. The dependence of these vectors on non-covalent molecular interactions demands high affinity binding to prevent off target effects from virus not bound to the adaptor. A potential solution is expression of the adaptor from a virus which is ablated for all-natural tropism. However, this is yet to be attempted in the context of an adaptor: virus complex and presents manufacturing difficulties which are yet to be overcome. More problematic is providing evidence that the adaptor molecule itself does not have detrimental interactions. This creates severe regulatory hurdles, leading many to abandon adaptor molecules in favour of antibody mimetic proteins that do not require disulphide bonds, thus able to integrate directly into the viral capsid.

1.5.2 Affibodies and FGFR2

Affibodies are antibody mimetic proteins based on a stable 3-helix bundle scaffold protein derived from Staphylococcus bacteria Protein A. They are versatile owing to their small size (~6.5kDa), ability to reach picomolar affinities, and to correctly fold in the cytosol^{214,215}.

A proof of concept study using an affibody specific for antibody Fc domain was fused to the fiber protein, replacing the knob domain, in a proof of concept study²¹⁶. The authors demonstrated that the modified virus can infect modified 293 cells displaying the Fc on their cell surface, but not wildtype (WT), non-Fc displaying, cells. However, the retargeted virus has a lower infectivity than the WT HAdV-C5 and the antibody to which the affibody binds is not a genetically incorporated modification and therefore the progeny virions lack specificity.

This strategy was implemented against the well-known cancer marker HER2/neu (AKA ERBB2)²¹⁷. The Fib Δ CAR-HI-Link-ZHZH fiber chimera contained 22 shaft repeats, an Δ LT_{485,486} deletion to ablate CAR affinity, and a

head to tail dimer of the anti-HER2 affibody (ZH) in the fiber knob HI loop (seen in figure 1-6A,E), was integrated into the Ad5/EGD vector (EGD being an integrin binding ablation mutation described in a previous study of affibody candidate molecules)¹⁷⁹. The Ad5/EGD/Fib Δ CAR-HI-Link-ZHZH virus infected HER2 expressing SKBR-3 (breast carcinoma) and SKOV-3 (ovarian carcinoma) cells with higher efficiency than WT HAdV-C5 or non-integrin binding ablated chimeric vectors. A comparable study using different anti-HER2 affibodies designed with an N-terminal trimerization motif (foldon, from T4 fibrin) entirely replacing the knob domain showed similarly effective results²¹⁸.

The Ad5/EGD/Fib Δ CAR-HI-Link-ZHZH includes a $_{91}\text{KKTK}_{94} \rightarrow _{91}\text{RKSK}_{94}$ mutation in the fiber-shaft to ablate HSPG interactions. The virus was later renamed Ad-ZH/3²¹⁹. When tested in mice bearing HER2^{high} PC346C prostate cancer tumours the Ad-ZH/3 cohort had significantly prolonged survival compared to mock, but in mice, survival did not significantly improve vs Ad5WT. Analysis of the mouse tumours revealed that those from the HAdV-C5WT treated mice retained HER2 expression in >40% of tumour mass. Ad-ZH/3 treated tumours were negative for HER2. As the authors note, this suggests that the virus has infected and killed HER2 positive tumours but left the remainder of the tumour mass to grow, effectively selecting a treatment resistant population.

This outcome is unlikely to occur in such an obvious manner in an immune competent model. The viruses immune stimulatory effects (now accepted as a primary mode of action for oncolytic viruses) are likely to activate a T-cell responses against neighbouring cancer cells^{2,220-223}. Regardless, this result is a reminder of the danger of monotargeted therapies, something accounted for in a proof of concept study using two affibodies, with different targets, inserted into the fiber HI-loop to create a virus with dual specificity²²⁴.

The most recent example of a virus retargeted using an affibody is HAdV-D43 pseudotyped with another affibody-knob chimera against HER2. The use of adenovirus HAdV-D43 leverages the low levels of pre-existing immunity to the rare species D Adenovirus and lack of cross neutralisation by anti-HAdV-C5 nAbs²²⁵. Whilst the virus can efficiently transduce HER2⁺ cells, it is hampered by poor production titres due to inefficient incorporation of the chimeric fiber.

In a departure from the usual fiber-chimeras the integration of an anti-HER2 retargeting affibody to the C-terminus of the pIX protein, C-terminal of an engineered cathepsin cleavage site²²⁶. Integration of peptides at the pIX C-terminal has been shown to be compatible with viral assembly^{227–230}. However pIX activity is required for dissociation of the fiber from the capsid during endosomal escape, to facilitate efficient infection²³¹. The authors predicted that high affinity association between pIX fusion proteins and ligand may prevent endosomal escape²³². Cathepsin is naturally present inside endosomes²³³, so the cathepsin cleavage site N-terminal of the affibody frees the affibody. This strategy enabled transduction of SKOV-3 spheroid cultures with greater efficiency than either Ad5WT or the pIX fusion lacking the cathepsin site, presumably by enabling cleavage of the affibody post-endocytosis and thus dissociation from the endosome.

Overall, affibodies represent an attractive means of retargeting Adenovirus. However, the chimeric fibers are often result in poor production titres, a limitation likely to hamper their potential for clinical translation. It is notable that none of the affibody retargeted viruses have included Factor X binding ablation mutations, which can result in sequestration of the virus in the liver. The extent of off target effects in the liver cannot be determined due to the lack of *in vivo* imaging of the infected mice, except for the Ad43 vector, which is shown not to interact with FX naturally. Any future development of these viruses must address these concerns.

1.5.3 Designed ankyrin repeat proteins (DARPs)

Another class of antibody mimetic molecules are DARPins (Designed Ankyrin Repeat Proteins). Similarly to affibodies (and many modern antibodies) they are generated by library creation from a stable scaffold (in this case ankyrin proteins) and biopanning^{234,235}. Unlike affibodies, DARPins have a larger MW (~14-28 kDa) and are structurally distinct, but otherwise similar in terms of the engineering opportunities in the context of adenovirus²³⁶.

Despite the free N and C-termini of DARPins, there has been little development of them as adenovirus fusion proteins. While similar adaptor strategies have been employed as with affibodies, enabling retargeting to H-Ras and HER2 (seen in Figure 6A,B,C), the DARPins are expressed separately, in *E.coli*, and

conjugated to the adenovirus prior to application^{237–239}. The lack of genetic incorporation of DARPins leaves them open to the same criticism as the previously mentioned anti-Fc domain affibody strategy. The lack of genetic incorporation prevents the passing of the targeting molecule on to progeny virions. This limits any potential therapeutic as they lose their ability to seek out the disease site after a single round of infection. However, these widely engineered molecules present a clear opportunity for genetic incorporation so are examined here.

A dimerised DARPin molecule was generated with specificity to both HER2 and adenovirus knob domain with an N-terminal SHP trimerization motif. This molecule has been reported to retarget HAdV-C5 to HER-2²³⁹. The authors used this in concert with a novel antibody derived construct, an scFv (mAb 9C12) against HAdV-C5 hexon trimerized with an N-terminal SHP motif. The authors demonstrated that the trimerized ScFv has pico-molar affinity for HAdV-C5 (K_D 10.4 pM). When complexed to the HAdV-C5 virus the DARPin construct efficiently retargeted the virus to HER2⁺ tumour cells, while the scFv shielded the virus from neutralising antibody activity¹⁶⁶.

This impressive feat of protein engineering presents a clear opportunity to produce the DARPin retargeting molecule *in cis* with the virus, either as a transgene or fusion protein to the fiber or pIX protein. Similarly, the efficacy of the shield adaptor may well be achieved by modification of the hexon hypervariable loops to eliminate immunogenic epitopes.

1.5.4 scTCR (single chain T-Cell Receptor) chimeric fiber proteins

An innovative study by Sebestyen *et al* used a virus expressing a chimeric fiber protein with an scTCR, specific to melanoma associated antigen A1 (MAGE-A1): Ad5.R1-scTCR, as a targeting protein, rather than an antibody mimetic molecule²⁴⁰. The fiber consisted of the N-terminal region of the HAdV-C5 fiber-shaft, the first pseudorepeat of the shaft, followed by an NRP trimerization motif and scTCR specific to HLA-A1 presented MAGE-A1 antigen. The virus initiated productive infection in an epitope specific manner, displaying efficient infection and transgene expression in cells expressing a MAGE-A1 epitope, but not MAGE-A2 or MAGE-A1 negative cell lines.

The melanoma associated cancer-testis antigen MAGE-A1 is one of many MAGE cancer antigens. The upregulation of MAGE's on cancer cells, which are normally restricted to germ cells, makes them an attractive target for therapeutic development owing to the potential for cancer selectivity^{241,242}.

While an approach dependent upon antigen recognition by TCRs is inherently restricted, both by patient HLA type and the availability of high affinity scTCRs, this presents an effective method of targeting cancer neo-antigens with virotherapy. As such it is intriguing that this is the first and last reported usage of a TCR retargeted adenovirus, such a vector could surely benefit from modern capsid detargeting mutations. Especially in light of more recent work demonstrating the potential effectiveness of oncolytic adenoviruses targeted to MAGE-A1, utilising non-genetic targeting technique (electrostatic coating of HAdV-C5 based conditionally replicating vector with MAGE-A1 peptide), and the ongoing development of therapeutic TCR molecules^{172,240,243}.

1.5.5 Retargeting by peptide incorporation

There have been many attempts to retarget adenovirus by integration of peptides specific to ligands of interest. Peptide insertion into the fiber knob can exceed 100 residues, 50% larger than the knob itself, without major detriment²⁴⁴. Though this is probably very dependent upon the character of the inserted sequence.

Attempts so far include incorporating peptides into the adenovirus fiber knob to target EGFR or FGFR1²⁴⁵. "Deknobbing" of the virus by removal of the fiber knob domain, and replacement with a trimerization motif and integrin binding RGD sequence²⁴⁶. Directed evolution approaches displaying peptides on the C-terminal of the fiber knob to generate affinity to a glioma, pancreatic cancer, Transferrin receptor, and thyroid carcinoma²⁴⁷⁻²⁵⁰. Further specific examples of peptide incorporation and a discussion of non-fiber knob incorporation sites including the pIX C-terminus are given in the excellent review by Dmitriev et al²⁵¹.

1.5.6 The A20 peptide to target $\alpha v \beta 6$

$\alpha v \beta 6$ integrin, an oncofoetal antigen, is highly expressed in many aggressive and invasive cancers but absent in healthy adult tissues making it an ideal

therapeutic target^{252,253}. Foot and Mouth Disease Virus (FMDV) naturally uses $\alpha\beta 6$ as a receptor, attaching via a semi-helical loop^{254,255}. It is from this interface that the A20FMDV2 peptide (A20) was derived, a 20mer with the amino acid sequence NAVPNLRGDLQVLAQKVART, shown to specifically bind $\alpha\beta 6$ with high affinity^{256,257}. A20 was integrated into the HI loop of HAdV-C5 fiber knob by genetic modification and shown to enable efficient infection of $\alpha\beta 6$ positive cells in a CAR independent manner²⁵⁸.

The addition of the TAYT mutation in order to ablate the native HAdV-C5 CAR tropism, which is not abolished by the A20 peptide, reduced hepatocyte tropism and uptake by liver resident macrophages (Kupffer cells)²⁵⁹. However, this was not enough to enable the virus to permeate the tumour mass in tested mice. The A20 peptide was incorporated into the fiber knob loops (CD and IJ) of an adenovirus with low seroprevalence, HAdV-D48, replicating the success of the Ad5.A20 vector. The HAdV-D48 fiber knob, in contrast to HAdV-C5, was shown to be insensitive to neutralisation by serum, raising hopes of a long sought after goal of oncolytic virotherapy: systemic delivery²⁶⁰.

Several HAdV-C5 based viruses were developed with and without the HAdV-D48 fiber knob pseudotype, or the KO1 (CAR binding ablation) mutation: Ad5.HI.A20, Ad5.KO1.HI.A20, Ad5/48kn.DG.A20. Respectively, these viruses were shown to transduce $\alpha\beta 6^+$ primary epithelial ovarian cancer (EOC) cells with 70, 60, and 16-fold improvements in affinity compared to HAdV-C5, and 160, 270, and 180-fold improved affinity in $\alpha\beta 6^{\text{high}}/\text{CAR}^{\text{low}}$ BT-20 triple negative breast cancer cells. Importantly, these modified A20 vectors appeared capable of infecting EOC cells in the presence of HAdV-C5 neutralising antibodies, demonstrating ≈ 1000 fold improvement in the transduction of EOC cells after pre-incubation with adenovirus neutralising ovarian ascites samples⁷⁹.

One example of the successful implementation of the A20 peptide retargeting technique is in the Ad5-3 Δ -A20T virus, in pancreatic cancer models. Three early phase viral genes were deleted (3 Δ) in this virus: E3gp19K to promote antigen presentation, E1ACR2 to render the virus conditionally replicative in tumour cells by preventing S-phase transition in non-replicating cells, and E1B19K to prevent apoptosis from being inhibited²⁶¹. The A20 FMDV2 peptide was inserted into the HAdV-C5 fiber knob protein, which was further modified with the TAYT

CAR binding ablation mutation (A20T). This is one of the most intricately engineered adenoviruses described to date, combining manipulation of the viral replication cycle via manipulation of the early genes to enhance cancer specific viral replication, with capsid modifications to improve viral specificity to tumour targets. The Ad5-3Δ-A20T was shown to kill effectively in cocultured pancreatic cancer and stromal cells, and prolong survival in xenograft models of pancreatic cancer in mice⁸⁰.

The latest example of the A20 peptide combined with adenovirus is the Ad5_{NULL}.A20 vector. While Ad5-3Δ-A20T relies upon early gene manipulation and preferential entry to the tumour cells, Ad5_{NULL}.A20 uses capsid modifications to achieve similar ends. In Ad5_{NULL}.A20 a KO1 mutation is used to ablate CAR interaction with the HAdV-C5 fiber knob and the A20 peptide is integrated into the HI-loop. This is combined with an RGE mutation in the penton base to prevent interaction with integrins, and an HVR7 mutation to inhibit binding to coagulation factor X. This virus showed a remarkable 10⁷-fold decrease in off target transduction of the liver, and 100%-survival at 100 days in a xenograft model of ovarian cancer in mice²⁶². The natural extension of the aforementioned technologies, Ad5-3Δ-A20T and Ad5_{NULL}.A20, is to combine them. This would theoretically create a virus capable of tumour specific transduction of tumour cells after systemic administration and tumour specific replication.

Despite successes, peptide retargeting is still limited by two key factors, which remain unaddressed: the availability of high affinity peptides for relevant tumour specific biomarkers, and inability to post-translationally modify adenovirus capsid proteins. While modern biopanning techniques can raise suitable peptides they are constrained by their nature as (usually) short 1-dimensional molecules, a structural constraint which frequently fails to translate into effective retargeting once integrated into a 3-dimensional viral protein.

This places peptides at a disadvantage compared to non-linear recognition motifs capable of forming 3-dimensional binding structures on antibodies and their mimetics^{263,264}. So far the only approaches to attempt to overcome this weakness in peptide design come in the form fiber knob chimeras utilising large, non-adenoviral proteins such as single-chain antibodies. It is possible that

modern *de novo* protein design techniques, in combination with biopanning, can be applied to the loop structure projecting from the adenovirus fiber knob to enable construction of highly avid 3-dimensional recognition sites against a pre-determined recognition epitope²⁶⁵.

The final constraint on peptide-based approaches to adenoviral retargeting is the lack of availability of PTMs. As the adenovirus capsid is assembled in the cytoplasm without trafficking through the ER or Golgi, precluding disulphide bond formation and glycosylation. Unless a method is found to facilitate *in vivo* post translational modification of adenoviral proteins without inhibiting virion assembly the peptide-based approach will remain limited to motifs which do not require PTMs.

1.6 Making adenovirus replication conditional upon oncogenic mutations in the cellular environment

As mentioned with the Ad5-3Δ-A20T virus, targeting strategies are not limited to capsid modification strategies⁸⁰. Modification of adenoviruses native replication strategy can make the virus conditionally replicating within cells with certain characteristics. The goal of post-entry modification strategies to make the adenovirus capable of replicating only within tumour cells, regardless of ingress method, thereby limiting the potential damage from on off-target infection of a non-cancerous cell. This can involve modification or deletion of the virus's native genes, or addition of non-native genetic elements.

The first oncolytic virus in clinical trial, called dl1520 then Onyx-015, was deleted for the E1B55K protein²⁶⁶⁻²⁷¹. E1B55K enables adenoviruses to replicate within cell which have functioning p53 protein. p53 arrests cell cycle growth in the presence of foreign (e.g. viral) DNA and in cells with damaged genomes. Adenovirus protein E1B55K binds to and inhibits the action of p53 to enhance G1/S cell cycle transition, consequently upregulating DNA and protein synthesis machinery to stimulate viral replication²⁷².

This makes E1B55K vital for adenovirus replication within normal cell. However, p53 is dysfunctional in around 50% of cancer, depending on type^{273,274}, making expression of E1B55K redundant in these cancers. Therefore, in theory,

deletion of the E1B55K gene from the viral genome renders the virus selectively replicating within p53 dysfunctional cancers.

In practice Onyx-105 was observed to replicate preferentially within cancer cell, but not due to a dependency upon p53 dysfunction²⁷⁵. Clinical trials yielded disappointing results for Onyx-015 and the next generation H101 virus, which is deleted for both E1B55K and E3B, which encodes immunomodulatory genes suppressing expression of MHC-I^{276,277}. It appears that mRNA export of virus mRNA from the cell nucleus was inhibited due to other E1B55K functions, binding to E4orf6 and suppression of host cell protein synthesis²⁷⁸. This leads to poor replication of Onyx-015 in cancer cell and associated low levels of cancer cell destruction and poor clinical outcomes²⁷⁸. However, this did not prevent approval of the H101 virus by the Chinese state food and drug administration (SFDA) as Oncorine²⁷⁹.

Engineering of early phase adenovirus genes therefore progressed to the E1A region, to avoid the issues associated with E1B55K deletion. The E1A gene is a requirement for adenovirus replication, it consists of conserved regions 1-4 (CR1-4). E1ACR2 binds to retinoblastoma protein (pRb), competitively inhibiting the transcription factor E2F from binding. This free E2F induces cell cycle entry into S-phase, again upregulating cellular machinery related to protein synthesis to aid viral replication²⁸⁰. Therefore adenoviruses with mutant E1ACR2 domains, such as dl922-947 (also called Ad Δ 24), are only capable of replicating in cancerous cells with a dysfunctional pRb pathway^{281,282}. The Ad Δ 24-RGD (DNX-2401) virus, retargeted to integrins, was shown to prolong glioma survival by 20% in a Phase-I clinical trial²⁸³. An HAdV-C5/B3 chimeric adenovirus with the Δ 24 mutation, and expressing granulocyte-macrophage colony stimulating factor (GM-CSF) as a transgene, has progressed to Phase-I and II clinical trials, in combination with chemotherapy under the name Oncos-102^{115-119,284}.

Another E1ACR2 deleted adenovirus, ICOVIR-7, controls E1A expression under an E2F promoter to increase expression in cells over-expressing this replication promoting gene²⁸⁵. There have been numerous iterations of this virus, including VCN-01 with the KKTK ablating RGDK motif to reduce HSPG interaction, which is in clinical trials to treat pancreatic ductal adenocarcinoma (PDAC)²⁸⁶⁻²⁸⁸.

The E3gp19K protein suppresses MHC-I antigen presentation on the cell surface²⁸⁹. Whilst early efforts in oncolytic virotherapy worked on the assumption that it was better to reduce immune recognition of the viral infection to facilitate viral replication, it is increasingly clear that immunogenicity of the virus helps recruit T-cells to the tumour and aid in tumour destruction²²³. A single base mutation within the E3gp19K gene (K455A) results in constitutively expressed membrane protein (E3gp19K-T1), by destroying the endoplasmic reticulum retention sequence. This was incorporated into the E1ACR2 deleted RGD-integrin targeted virus ORCA-010²⁹⁰. This E3gp19K-T1 gene results in permeabilization of the cell membrane, assisting with progeny virion release and viral spread.

Another virus, ColoAd-1 or Enadenotucirev (EnAd) was not generated by rational design, but instead used a selected evolution method⁹⁷. This virus was generated by facilitating recombination between various species of adenoviruses in HT29 colon cancer cells and selecting for the virus with the greatest rate of replication. EnAd is predominantly a chimera of species B adenoviruses HAdV-B3 and HAdV-B11p, but has lost the entire E3-region and the E4orf4. The rarer HAdV-B11p capsid was less efficiently neutralised by human serum than HAdV-C5 which has relatively high seroprevalence⁹⁹. E4orf4 limits E1A expression and helps to control AMPK related cell metabolism²⁹¹.

The various replication controlling genetic mutants of adenovirus have been regulated in many ways, including the use of tumour specific promoters^{292–294}. Predominantly replication has been dependent upon controlling E1A expression. An example of this is the AdV-TERTp-E1A virus which regulates E1A expression with the human telomerase reverse transcriptase (hTERT) promoter, making expression of this critical adenovirus replication protein commensurate with that of the hTERT proteins which is often upregulated in cancer²⁹⁵.

Human prostate specific antigen (PSA) is often used as a diagnostic marker for prostate cancer as it is upregulated. The CN706 is another example of selective expression of E1A, in this case controlled by a promoter construct derived from the 5' flanking region of the PSA gene. When developed in the 1990's this virus showed promise but did not progress to an approved therapy²⁹³. More recently,

the CG7870 virus also targets prostate cancer, controlling E1A and E1B expression with rat probasin promotor and PSA promoter, respectively^{296,297}. This virus progressed to clinical trial and it's double promotor control serves to enhance specificity further, requiring both tumour associated genes to be sufficiently upregulated to facilitate viral replication.

The tumour specific promotor method of controlling viral replication is still being refined. Since human genes are usually expressed in healthy human tissues as well as cancerous ones, promotor based control can cause off target replication of the virus and, theoretically, damage to healthy tissues. This can be overcome by selecting genes which are upregulated specifically in cancer and/or adjusting promotor strength to only be effective at very high levels of gene expression associated with cancer. However, overly tight regulation will hamper viral replication and therefore oncolysis, a major source of the therapeutic activity. The nature of this balance will depend upon the tumour cells targeted and the specificity, packaging efficiency, and replication rate of the virus.

1.7 Clinical trials of adenovirus derived oncolytic viruses

As of 31/07/2019 there are 89 active clinical trials studying adenovirus-based drugs (fig1-7). Of these four are in Phase-III clinical trials, though 3 of these are, arguably, not oncolytic virotherapies but cancer gene therapies. In these trials of INSTILADRIN (rAd-IFN)^{298,299} and ADV-Tk³⁰⁰ the adenovirus is being used a non-replicating gene transfer vector for interferon α and herpes simplex thymidine kinase. The goal is to sensitise cancer cells to drug based treatment rather than use the virus itself as a means of destroying the cancer³⁰¹⁻³⁰³.

The oncolytic adenovirus in Phase-III clinical trial is H101. This engineered adenovirus type 5 is E1B55K and E3 deleted, designed to make replication p53 dependent and present viral antigens via MHC. In this trial H101 is being trailed in combination with hepatic artery infusion chemotherapy (HAIC). No results are available yet³⁰⁴.

A subset of clinical trials which are active at the time of writing is shown in table 2. The majority of these are in phase-I, examining safety. Also, a majority are in combination with chemotherapy, highlighting the fact that whilst oncolytic adenoviruses are presently effective as combination therapies, they are not yet

viable agents alone³⁰⁵. However, LOAd-703 is showing promise, having progressed to phase-I/II clinical trials as a single agent. This adenovirus is a CD46 receptor binding HAdV-C5 virus with the fiber protein of HAdV-B35, with

Table 1-2: A subset of active clinical trials utilising adenovirus derived oncolytic viruses.

NCT Number	Title	Conditions	Interventions	Phases	First Posted
NCT03916510	Chemoradiation With Enadenotucirev as a Radiosensitiser in Locally Advanced Rectal Cancer	Locally Advanced Rectal Cancer	Biological: Enadenotucirev Drug: Capecitabine Radiation: Radiotherapy	Phase 1	16/04/2019
NCT03896568	Oncolytic Adenovirus DNX-2401 in Treating Patients With Recurrent High-Grade Glioma	IDH1 wt Allele Recurrent Anaplastic Astrocytoma Recurrent Glioblastoma Recurrent Gliosarcoma Recurrent Malignant Glioma	Biological: Oncolytic Adenovirus Ad5-DNX-2401 Procedure: Therapeutic Conventional Surgery	Phase 1	01/04/2019
NCT03852511	First in Human Study of NG-350A (an Oncolytic Adenoviral Vector Which Expresses an Anti-CD40 Antibody)	Metastatic Cancer Epithelial Tumor	Biological: NG-350A	Phase 1	25/02/2019
NCT03773744	MG1-MAGEA3 With Ad-MAGEA3 and Pembrolizumab in Patients With Previously Treated Metastatic Melanoma or Cutaneous Squamous Cell Carcinoma	Metastatic Melanoma Squamous Cell Skin Carcinoma	Drug: Ad-MAGEA3 Drug: MG1-MAGEA3 Drug: Pembrolizumab Drug: Cyclophosphamide	Phase 1	12/12/2018
NCT03740256	Binary Oncolytic Adenovirus in Combination With HER2-Specific CAR VST, Advanced HER2 Positive Solid Tumors (VISTA)	Bladder Cancer Head and Neck Squamous Cell Carcinoma Cancer of the Salivary Gland Lung Cancer Breast Cancer Gastric Cancer Esophageal Cancer Colorectal Cancer Pancreatic Adenocarcinoma	Biological: CAdVEC	Phase 1	14/11/2018
NCT03714334	DNX-2440 Oncolytic Adenovirus for Recurrent Glioblastoma	Glioblastoma Glioblastoma, Adult	Drug: DNX-2440 injection	Phase 1	22/10/2018
NCT03618953	This is a Trial of MG1-E6E7 With Ad-E6E7 and Atezolizumab in Patients With HPV Associated Cancers	HPV-Associated Cancers	Biological: Ad-E6E7 Biological: MG1-E6E7 Biological: Atezolizumab	Phase 1	07/08/2018
NCT03281382	Phase 1 Trial of Interleukin 12 Gene Therapy for Metastatic Pancreatic Cancer	Metastatic Pancreatic Cancer	Biological: Ad5-yCD/mutTKSR39rep-hIL12	Phase 1	13/09/2017
NCT03225989	Trial Investigating an Immunostimulatory Oncolytic Adenovirus for Cancer	Pancreatic Adenocarcinoma Ovarian Cancer Biliary Carcinoma Colorectal Cancer	Drug: LOAd703	Phase 1 Phase 2	21/07/2017
NCT03190824	Evaluate Efficacy, Immunological Response of Intratumoral/Intralesional Oncolytic Virus (OBP-301) in Metastatic Melanoma	Melanoma Stage III Melanoma Stage Iv	Drug: OBP-301	Phase 2	19/06/2017
NCT03178032	Oncolytic Adenovirus, DNX-2401, for Naive Diffuse Intrinsic Pontine Gliomas	Brainstem Glioma Neoadjuvant Therapy	Biological: DNX-2401	Phase 1	06/06/2017
NCT03072134	Neural Stem Cell Based Virotherapy of Newly Diagnosed Malignant Glioma	Glioma Anaplastic Astrocytoma Anaplastic Oligodendroglioma Anaplastic Oligoastrocytoma Glioblastoma Multiforme Astrocytoma, Grade III Astrocytoma, Grade IV Brain Cancer	Biological: Neural stem cells loaded with an oncolytic adenovirus	Phase 1	07/03/2017
NCT03029871	Oncolytic Adenovirus-Mediated Gene Therapy for Lung Cancer	Non-small Cell Lung Cancer Stage I	Biological: Ad5-yCD/mutTKSR39rep-ADP Adenovirus	Phase 1	24/01/2017

NCT03004183	SBRT and Oncolytic Virus Therapy Before Pembrolizumab for Metastatic TNBC and NSCLC	Metastatic Non-small Cell Lung Cancer Metastatic Triple-negative Breast Cancer	Biological: ADV/HSV-tk Drug: Valacyclovir Radiation: SBRT Drug: Pembrolizumab	Phase 2 28/12/2016
NCT03003676	A Pilot Study of Sequential ONCOS-102, an Engineered Oncolytic Adenovirus Expressing GMCSF, and Pembrolizumab in Patients With Advanced or Unresectable Melanoma Progressing After Programmed Cell Death Protein 1 (PD1) Blockade	Advanced or Unresectable Melanoma Progressing After PD1 Blockade	Biological: ONCOS-102 Drug: Cyclophosphamide Drug: Pembrolizumab	Phase 1 28/12/2016
NCT02879760	Oncolytic MG1-MAGEA3 With Ad-MAGEA3 Vaccine in Combination With Pembrolizumab for Non-Small Cell Lung Cancer Patients	Non-Small Cell Lung Cancer	Biological: Ad-MAGEA3 Biological: MG1-MAGEA3 Biological: Pembrolizumab	Phase 1 Phase 26/08/2016
NCT02879669	A Randomised Phase II Open-label Study With a Phase Ib Safety lead-in Cohort of ONCOS-102, an Immune-priming GM-CSF Coding Oncolytic Adenovirus, and Pemetrexed/Cisplatin in Patients With Unresectable Malignant Pleural Mesothelioma	To Determine Safety, Tolerability and Efficacy of ONCOS-102 in Combination With Chemotherapy	Biological: ONCOS-102 Drug: Pemetrexed/cisplatin (carboplatin) Drug: Cyclophosphamide	Phase 1 Phase 26/08/2016
NCT02798406	Combination Adenovirus + Pembrolizumab to Trigger Immune Virus Effects	Brain Cancer Brain Neoplasm Glioma Glioblastoma Gliosarcoma Malignant Brain Tumor Neoplasm, Neuroepithelial Neuroectodermal Tumors Neoplasm by Histologic Type Neoplasm, Nerve Tissue Nervous System Diseases	Biological: DNX-2401 Biological: pembrolizumab	Phase 2 14/06/2016
NCT02705196	LOAd703 Oncolytic Virus Therapy for Pancreatic Cancer	Pancreatic Cancer	Biological: LOAd703 Drug: Gemcitabine Drug: nab-paclitaxel	Phase 1 Phase 2 10/03/2016
NCT02555397	Phase 1 Trial of Interleukin 12 Gene Therapy for Locally Recurrent Prostate Cancer	Prostate Cancer	Biological: Ad5-yCD/mutTKSR39rep-hIL12	Phase 1 21/09/2015
NCT02285816	MG1 Maraba/MAGE-A3, With and Without Adenovirus Vaccine, With Transgenic MAGE-A3 Insertion in Patients With Incurable MAGE-A3-Expressing Solid Tumours	Advanced/Metastatic Solid Tumours	Biological: MG1MA3 Biological: AdMA3	Phase 1 Phase 2 07/11/2014
NCT02045602	Phase I Dose Escalation Study of Intravenous VCN-01 With or Without Gemcitabine and Abraxane in Patients With Advanced Solid Tumors	Locally Advanced Solid Tumors Metastatic Solid Tumors Pancreatic Adenocarcinoma	Genetic: VCN-01 Drug: Gemcitabine Drug: Abraxane®	Phase 1 27/01/2014

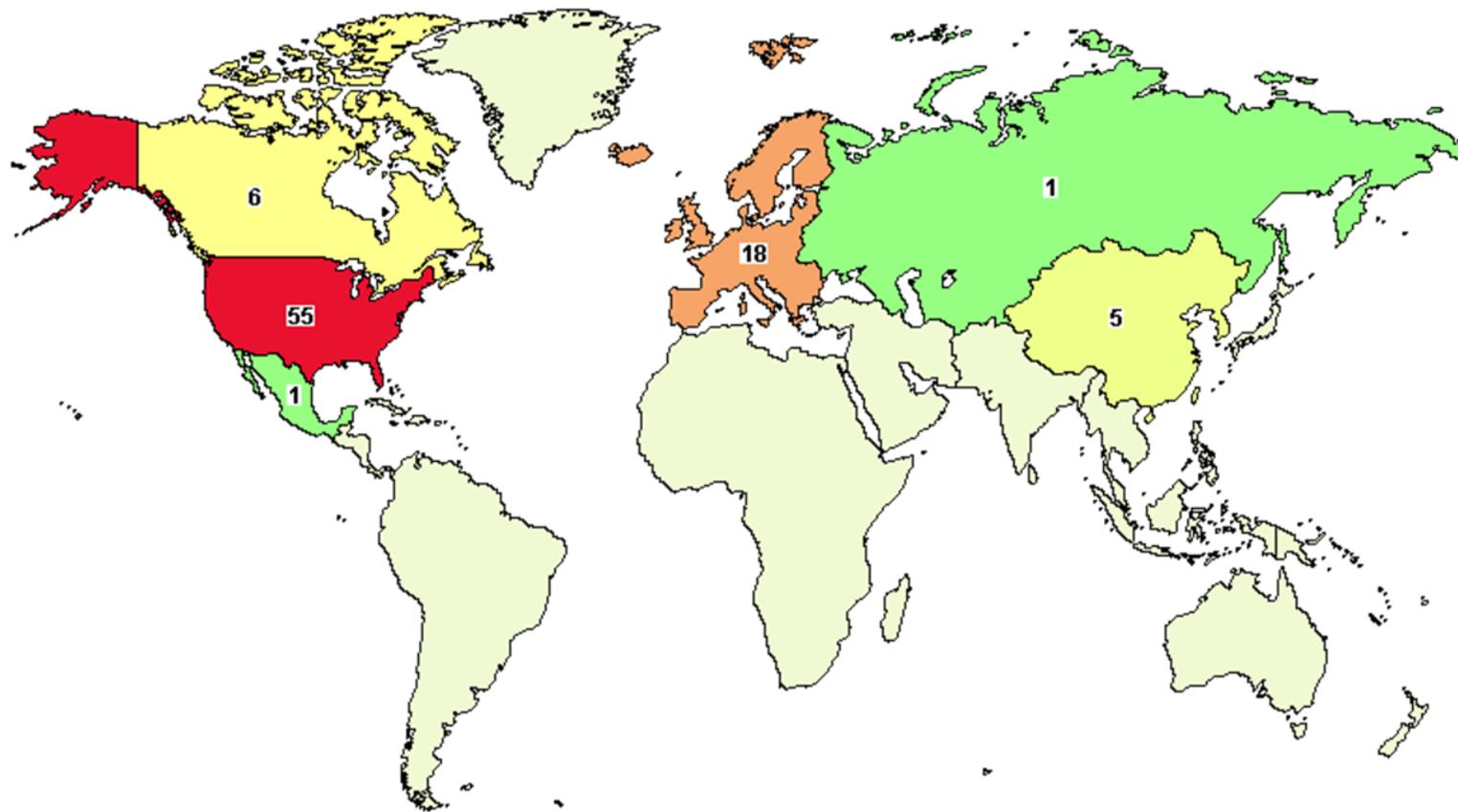
promoter controlled E1A expression, engineered to express CD40L and 4-1BB immunomodulators^{306–308}.

It is notable that none of the clinical candidates discussed so far are targeted via their capsid. Whilst LOAd-703 and Enadenotucirev are engineered, they both use CD46 as a receptor, which is ubiquitously expressed on human cell surfaces. This shows that, from a safety perspective, it is enough to control adenovirus replication through a combination of careful administration, dose control, and regulated replication post-infection. However, these methods also have associated disadvantages.

If an oncolytic virus interacts with, and infects, cells non-specifically then it will enter many cells which are not the tumour. This is often overcome by intratumoural administration of the oncolytic virus directly into the tumour. However, this method negates a key advantage of oncolytic virotherapy: systemic dissemination of the virotherapy to tackle metastatic tumour sites. Instead, any progeny virions made within the tumour will spread locally. Initially this will be into the tumour mass, but as the tumour mass decreases progeny virions will be absorbed by surrounding cells expressing the receptor for the virotherapy and leakage of the therapeutic vector away from the tumour. These off-target interactions also mean that for any given number of progeny, more will enter cells where they have no therapeutic benefit, reducing the overall efficacy of a given therapeutic dose and requiring a larger number of viral particles to be efficacious. These issues can only be resolved at the primary point of infection, which is cellular attachment and internalisation.

The adenovirus targeting methods discussed so far all have drawbacks. These are associated with difficulty in achieving specificity, or in damaging the viruses efficacy as a drug. A multilevel targeting approach is likely to be the key to overcoming these challenges.

LOAd-703 has attempted multilevel targeting, using a combination of two different tumour specific promoters. While this may alleviate issues with off-target replication, it does not relieve problems with off-target infection. The Ad5_{NULL}.A20 virus achieves exquisite levels of tumour selectivity at the infection level but is limited to cancers with the $\alpha\beta 6$ surface receptor.



Colors indicate the number of studies with locations in that region.

Least  Most

Labels give the exact number of studies.

Figure 1-7 Map of the number and locations of active clinical trials using adenovirus vectors from clinicaltrials.gov using the following search criteria: on 30/07/2019: Recruiting, Not yet recruiting, Available, Active, not recruiting, Enrolling by invitation, Approved for marketing Studies | Cancer | Adenovirus OR AdV OR Ad | Phase Early Phase 1, 1, 2, 3, 4

Huge progress has been made in achieving tumour specific infection with adenovirus based oncolytic viruses. However, no virus has yet attempted to overcome the next inevitable problem in the field, which is tumour heterogeneity. Every virus discussed in this introduction and, to the best of our knowledge, under development, is directed against a single target. Whether this target is a cell surface receptor or a mutated intracellular pathway, the fact remains that the virotherapies mode of action is dependent upon a particular cancer phenotype. Tumour heterogeneity is now an accepted part of cancer biology, and plays a significant role in recurrence after remission³⁰⁹. Oncolytic viruses have an inherent advantage in respect to attacking heterogeneous tumour cells, as they engage the immune system which can then recognise and attack other tumour cells not directly affected by the virus. However, it remains to be seen if this is efficient and could well be improved by making the virus itself capable of attacking multiple tumour targets.

1.8 Aims

The aims of this study were as follows:

1. **Identify phenotypes potentially useful for the treatment of cancer in adenovirus species D**
2. **Clarify the primary cell entry receptors being utilised by adenovirus species D**
3. **Identify previously unknown receptor usage by adenovirus species D**
4. **Understand the mechanisms which govern adenovirus species D infection**

The work in this thesis aims to enhance our understanding of adenovirus mechanisms of infection and receptor tropism in order to engineer more effective viral vectors. We begin by investigating an understudied species of adenoviruses, species D, and observing a type with potentially beneficial infection properties for gene transfer, and as a cancer vaccine. We continue our investigation of the species D adenoviruses, seeking to clarify a long-standing controversy within the field regarding the usage of CD46 as a primary receptor tropism. We clarified this interaction and demonstrate affinity for a different receptor: CAR, in the processes showing that fiber knob hypervariable loops play a role in regulating fiber knob proteins through steric hindrance⁷². We continued our investigation into receptor tropisms of species D adenoviruses and demonstrated a previously unknown tropism towards sialic acid for HAdV-D26K. HAdV-D26 is a clinically important adenovirus due to its usage as a vaccine vector in phase-III clinical trials. Our investigation of HAdV-D26 demonstrates that sialic acid binding is not restricted to adenoviruses which cause epidemic keratoconjunctivitis (EKC) as has previously been thought⁵⁸. Instead, the results presented here-in show that several other species D adenoviruses with no know EKC association can interact with sialic acid and that it may be a broadly utilised receptor for the entire adenovirus species.

CHAPTER 2. Materials & Methods

2.1 Sequence comparison and Phylogenetic Tree generation

2.1.1 Sequence alignment and visualisation

All sequence alignments were performed using the EMBL-EBI Clustal Omega tool^{310,311}. Sequences were visualised using BioEdit (V5.10, Isis Therapeutics). Fiber knob sequences were defined as the conserved TLW hinge motif to the C-terminus of the protein.

2.1.2 Wu-Kabat variability analysis

Amino acid sequence variability scores were calculated from the Clustal omega aligned amino acid sequences of the fiber knob domains of either adenoviruses 1-56, or only the species D adenoviruses. Analysis was performed using the protein variability server (PVS), using a consensus base sequence and the Wu-Kabat method³¹².

2.1.3 Phylogenetic Maximum likelihood tree generation

Representative whole genomes (nucleotide) of adenoviral species 1-56 were selected from the National Center for Biotechnology Information (NCBI), and aligned using the EMBL-EBI Clustal Omega tool^{310,311}. Fiber knob domain amino acid sequences of the fiber knob proteins were derived from the same genome sequences, defined as the translated nucleotide sequence of the fiber protein (pIV) from the conserved TLW hinge motif to the protein C-terminus, and aligned in the same manner as the whole genomes. Phylogenetic relationships were inferred using the maximum likelihood method based upon the Jukes Cantor model for the whole genome nucleotide analysis³¹³, and the Poisson correction model for the fiber knob amino acid analysis³¹⁴, using MEGA X software³¹⁵. Confidence was determined by bootstrap analysis (500 replicates)³¹⁶ and trees displayed condensed at 70% confidence (percentage confidence values shown at each node).

Genomic sequences from which fiber knob domain sequences were determined, which have been used in phylogenetic analysis, have the following NCBI Accession numbers: AC_000017| AF532578| X73487| AY803294| AB562586| AY601636| AF108105| GU191019| JQ326209| AC_000007| JN226749| KF528688| FJ404771| JN226750| JN226751| JN226752|

EF153474|JN226753| FJ824826| JN226754| JN226755| AM749299| JN226756|
JN226758| AY737797| AC_000019| GQ384080| JN226759| JN226760|
KU162869| DQ315364| JN226761| JN226762| JN226763| JN226764|
AY875648| JN226757| EF153473| DQ393829| AC_000008| AY737798|
JN226765| DQ923122| AB605243| NC_012959| FJ643676|HM770721|
HQ413315| AC_000018| DQ086466| JN226746| JN226747| AB448776|
AB448767| AJ854486| KF006344|

2.2 Expression of recombinant fiber knob proteins

SG13009 *E.coli* harbouring pREP-4 plasmid and pQE-30 expression vector containing the fiber knob DNA sequence (synthesised by TWIST Biosciences) were cultured in 20ml LB broth with 100µg/ml ampicillin and 50µg/ml kanamycin overnight from glycerol stocks made in previous studies^{79,258,260}. 1L of TB (Terrific Broth, modified, Sigma-Aldrich) containing 100µg/ml ampicillin and 50µg/ml of kanamycin were inoculated with the overnight *E.coli* culture and incubated at 37°C until they reached an optical density (OD) of 0.6 at λ570nm. 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).

2.3 Cell Culture Methods

2.3.1 Culture of cell lines

Cell lines were maintained at 20-90% confluency in T150 cell culture flasks (Corning) in a human tissue act (HTA) certified cell culture incubator (HERA Cell, Thermofisher) at 37°C under a 5% CO₂ humidified atmosphere. Media was changed under sterile conditions as necessary until cells reached 90% confluency.

Cells were sub-cultured every 3-4 days at a 1:5 ratio as follows. Cells were washed with 10ml phosphate buffered saline (PBS) and 5ml of 0.05% Trypsin-EDTA (Gibco) was applied. Cells were incubated until they lost adherence to the culture flask, the monolayer dissociated by pipetting, and cells passed through a strainer (Cole-Parmer). 1ml of trypsinised and strained cells were applied to a new T150 cell culture flask and 20ml of culture media (described in Table 2-1) added.

All cell lines were routinely tested for mycoplasma contamination and found to be negative.

Table 2-1 Cell lines used in this study, their origins, and the media which was used to support them.

Cell Line Name	Origin of cell line	Culture Media	FBS (%v/v)	Supplements
293 T-Rex	Human Embryonic Kidney	DMEM	10	5mM Glutamine
A427	Lung carcinoma	EMEM	10	5mM Glutamine
A549	Lung carcinoma	DMEM	10	5mM Glutamine
ASPC1	Pancreatic adenocarcinoma (Ascitic RPMI 1640 metastasis)		10	5mM Glutamine, 1mM NaP
BT-20	Breast carcinoma	DMEM	20	5mM Glutamine

BT-474	Breast ductal carcinoma	DMEM	20	5mM Glutamine
BxPc3	Pancreatic adenocarcinoma	RPMI 1640	10	5mM Glutamine
CFPAC-1	Pancreatic ductal adenocarcinoma (Liver metastasis)	DMEM	10	5mM Glutamine
CHO-BC1	Chinese Hamster Ovary	DMEM-F12	10	5mM Glutamine
CHO-CAR	Chinese Hamster Ovary	DMEM-F12	10	5mM Glutamine
CHO-K1	Chinese Hamster Ovary	DMEM-F12	10	5mM Glutamine
DLD-1	Colorectal adenocarcinoma	DMEM	10	5mM Glutamine
hMDDC	Human monocyte derived dendritic cells (Human Blood)	IMDM	10	5 U/ml GM-CSF, 5 U/ml IL-4, 50 μ M β 2-mercaptoethanol
Ky5e-30	Eosophageal squamous cell carcinoma	RPMI 1640 + Ham's F12 (1:1)	2	5mM Glutamine
MCF7	Breast adenocarcinoma (Pleural effusion metastasis)	EMEM	10	5mM Glutamine
MDA-MB-231	Breast adenocarcinoma (Pleural effusion metastasis)	DMEM	20	5mM Glutamine
MDA-MB-361	Breast carcinoma (Brain metastasis)	DMEM	20	5mM Glutamine

MIA PaCa2	Pancreatic adenocarcinoma	DMEM	10	5mM Glutamine
mMDDC	Mouse monocyte derived dendritic cells (Mouse Blood)	RPMI 1640	10	5U/ml GM-CSF, 5 U/ml IL-4
NCI-H460	Large cell lung carcinoma	RPMI 1640	10	5mM Glutamine
Panc0403	Pancreatic ductal adenocarcinoma	RPMI 1640	10	5mM Glutamine, 10mM HEPES, 1mM NaP
Panc10	Pancreatic adenocarcinoma	DMEM	10	5mM Glutamine
PT45	Pancreatic ductal adenocarcinoma	RPMI 1640	10	5mM Glutamine
SKOV-3	Ovarian adenocarcinoma (Acitic metastasis)	DMEM	10	5mM Glutamine
Suit2	Pancreatic carcinoma (Liver metastasis)	DMEM	10	5mM Glutamine
SW1990	Pancreatic adenocarcinoma (Splenic metastasis)	DMEM	10	5mM Glutamine

2.3.2 Long term cell storage

Cell lines were grown until fully confluent in T150 cell culture flasks (Corning), washed and trypsinised as described previously (section 2.3.1) then neutralised in complete media with a 1:1 ratio. Cells were then centrifuged for 5mins at 1200 rpm and the supernatant discarded. The cell pellet was then resuspended

in 3ml of Fetal bovine serum containing 10% (v/v) of dimethyl sulfoxide (DMSO). 1ml aliquots were dispensed into 1.5ml cryovials (Corning) and placed into Mr Frosty freezing containers (ThermoFisher) according to the manufacturer's instructions and frozen overnight at -80°C. Cryovials were then placed in vapour phase liquid nitrogen tanks for long term storage.

To resuscitate the cells they were thawed at 37°C. The cells were then pipetted into a T150 cell culture flask containing 20ml of cell culture media and incubated at 37°C overnight to allow the cells to adhere. The media was then replaced and cells cultured as normal.

2.3.3 Monocyte derived dendritic cells (MDDC)

MDDC were a kind gift from collaborators. James Wheeldon (Piquet group, Cardiff university) provided human MDDC, and Marta Williams (Humphreys group, Cardiff University) provided mouse MDDC.

2.4 Virus generation

2.4.1 Generation of pseudotype virus bacterial artificial chromosomes

Pre-existing bacterial artificial chromosomes (BACs) encoding the genome of replication deficient (E1A and E3 deleted) HAAdV-C5 virus expressing luciferase or GFP as a transgene were used as the template for the pseudotyped adenovirus genomes used in this study. The fiber knob is defined as beginning from the TLW hinge motif, which is conserved in all adenovirus fiber proteins¹⁴. The DNA sequences of the fiber knob proteins used in this study were synthesised commercially (GeneWiz, or TWIST Biosciences). The pre-existing BAC was modified using the AdZ recombineering methodology developed by Stanton *et al*³¹⁷. This methodology utilises homologous recombination in SW102 *E.coli* harbouring the BAC to integrate a marker cassette at the site to be engineered, and then a second round of homologous recombination to replace the marker cassette with the gene of interest; in this case the replacement fiber knob protein.

Primers were designed to have homology to a RPSL DNA selection cassette, provided by Dr Richard Stanton, at the 3' end³¹⁷. The 5' end of the primer was designed to have 70bp of homology to the region of interest. For the forward

primer this is the region directly 3' of the TLW hinge DNA sequence, and, for the reverse primer, the region immediately 5' of the end of the HAdV-C5 fiber open reading frame. These primers were used to amplify the RPSL cassette DNA by PCR (HiFi expand DNA polymerase system, ThermoFisher). The resultant fiber knob region specific DNA selection cassette was purified from the PCR mix using the QIAquick PCR Purification kit (QIAGEN).

The SW102 *E.coli* harbouring the HAdV-C5 BAC were cultured to OD 0.6 (600nm wavelength) in 25ml of LB at 32°C in a shaking incubator. The recombination genes were then induced by incubation at 42°C for 15mins. The SW102 were then cooled on ice for 20mins, then pelleted by centrifugation at 4000RPM for 5mins. The supernatant was discarded, and the pellet was washed by resuspension in 50ml of cold ddH₂O and centrifugation, twice. The pellet was then resuspended in residual ddH₂O after the supernatant was discarded. 50µl was transferred to pre-cooled electroporation cuvettes (2mm HiMaX, Isogen life science) and 4µl of purified fiber knob region specific RPSL DNA selection cassette was added and electroporated (Micropulser electroporator, BioRad) using pre-set program Ec3. 1ml of LB was then added to the electroporated SW102 and incubated at 32°C in a shaking incubator for 1hr to recover. 50µl of these cells were plated out onto LB agar (1.5%) containing 15µg/ml Kanamycin, 12.5µg/ml chloramphenicol, 25µg/ml of X-GAL, and 200mM IPTG and cultured at 32°C for 48hrs. A blue colony was then selected and inoculated into two different culture conditions: LB containing 15µg/ml Kanamycin and 12.5µg/ml chloramphenicol, or LB containing 400µg/ml streptomycin and 12.5µg/ml chloramphenicol. This was done to verify that the blue colony contained the gene encoding streptomycin sensitivity, so only colonies which failed to grow in streptomycin containing media were utilised in the next steps.

The synthesised DNA sequences encoding the fiber knob genes were amplified by PCR (HiFi expand DNA polymerase system, ThermoFisher) using primers with the same 3' homology sequences as used to integrate the marker cassette and a 5' sequence homologous to the fiber knob DNA. The resultant DNA, the fiber knob encoding sequence with homology arms to either side of the RPSL cassette, was purified using the QIAquick PCR Purification kit (QIAGEN).

The selected *E.coli* from the previous step which now contain the HAdV-C5 BAC with the RPSL marker in place of the fiber knob DNA were prepared for electroporation in an identical manner to the first step. 4µl of the fiber knob DNA with homology arms to the regions either side of the integrated RPSL cassette was then mixed with the cells in the cuvette, electroporated, and recovered as in the previous step. 50µl of these cells were then plated out onto an LB agar (1.5%) plate containing 400µg/ml streptomycin, 12.5µg/ml chloramphenicol, 25µg/ml of X-GAL, and 200mM IPTG and cultured at 32°C for 48hrs. A white colony was then selected from the plate and inoculated into 5ml LB containing 12.5µg/ml chloramphenicol, then cultured at 32°C overnight in a shaking incubator. This culture was used to generate a glycerol stock by mixing 300µl of glycerol with 700µl of culture in a cryovial and frozen at -80°C. DNA was extracted from the rest of the culture using a miniprep kit (QIAGEN). The purified BAC DNA was sequenced commercially (Eurofins) to determine the correct fiber knob sequence is present in the new BAC.

2.4.2 Amplification of BAC DNA

SW102 *E.coli* containing the correct BAC DNA encoding viral genomes were cultured at 32°C in 10ml of LB containing 12.5µg/ml chloramphenicol for 6-8hrs prior to inoculation of this starter culture into 250ml of LB containing 12.5µg/ml chloramphenicol and incubation at 32°C overnight. This culture was harvested by centrifugation at 6000RPM for 20min at 4°C in a high-speed centrifuge (Avanti J-20 XP, Beckman Coulter). This pellet was resuspended by pipetting in 24ml of buffer S1 (NucleoBond BAC 100 Kit, Machery-Nagel), then mixed with buffer S2 (NucleoBond BAC 100 Kit, Machery-Nagel) and mixed by inversion and incubated for 5min at room temperature. 24ml of buffer S3 (NucleoBond BAC 100 Kit, Machery-Nagel), pre-cooled on ice, was then added, mixed by inversion, and cooled on ice for 5mins. This mixture was then centrifuged at 6000RPM for 15min at 4°C, whilst the BAC purification column was equilibrated by addition of 6ml of buffer N2 (NucleoBond 100 Kit, Machery-Nagel). The supernatant was passed through the column by gravity, then the column was washed twice with 18ml of buffer N3 and the flow through discarded. The DNA was then eluted with 15ml of buffer N5 pre-heated to 50°C and the eluant mixed with 11ml of isopropanol. The solution was then centrifuged at 15000RPM for

30min at 4°C to pellet the DNA. The pellet was then washed with 70% ethanol (30% H₂O, v/v), vortexed, and pelleted again at 15000RPM for 30min at 4°C. The wash solution was then discarded, and the pellet allowed to dry at 37°C. The DNA pellet was then resuspended in 100µl of water and allowed to dissolve for 15mins. DNA was then stored at -20°C until ready to be used for virus transfection.

2.4.3 Transfection of BAC DNA and virus rescue

Transfections were performed in 293 T-Rex cells using Effectene transfection kit (QIAGEN). 5×10^5 293 T-Rex cells were seeded in a T25 CellBind cell culture flask (Corning) and allowed to adhere overnight in 5ml of complete cell culture media at 37°C in a humidified 5% CO₂ atmosphere. 1µg of BAC DNA was mixed with 8µl of buffer EC, mixed by pipetting, and incubated at room temperature for 5min. 25µl of Effectene transfection reagent was then added to the DNA enhancer mix, mixed by inversion, and incubated at room temperature for 5min. This mixture was added to 1ml of cell culture media and the whole mixture added to the prepared 293 T-Rex cell culture, drop-wise, and then replaced in the cell culture incubator. A second T25 flask was treated identically except for the absence of the BAC DNA to be used as a control. Media was removed from the transfected cell culture after an overnight incubation and replaced with 5ml of complete cell culture media. The cells were then cultured as normal, exchanging media when it became discoloured and inspecting the cell monolayer for signs of cytopathic effect (CPE, rounding of the cells and loss of adherence to the cell surface), which usually occurred in 7-14 days, and should be apparent in the BAC transfected cells and not the control flask. When CPE was observed the monolayer was closely monitored and allowed to progress through the entire monolayer. Cells exhibiting CPE were then dissociated from the cell surface by tapping the side of the culture flask. Cells were then collected in the cell culture media and centrifuged at 1200RPM for 5min in a 15ml falcon tube (Corning) to pellet them. 1ml of supernatant was retained and the rest discarded. The cell pellet was resuspended in the retained 1ml of cell culture media. This 1ml suspension of infected cells was transferred to a 1.5ml cryovial and stored at -80°C.

2.4.4 Propagation of rescued viruses

293 T-Rex cells were sub-cultured into 11x T150 CellBind cell culture flasks (Corning) and grown to approximately 80% confluence. Media was then replaced on one flask to be used as a control. 250ml of complete cell culture media was mixed with 20 μ l of the cell suspension from the rescue of the virus by transfection of the genome containing BACs. This infected media was used to replace the cell culture media on the 10 remaining T150 cell culture flasks. Flasks were cultured under normal 293 T-rex cell culture conditions, replacing the media as it became discoloured, and inspecting for CPE by comparison to the control flask. When CPE became apparent (usually 4-7 days) it was allowed to propagate through the entire monolayer and the cells were harvested by tapping on the flask to dislodge the cells, collection of the media containing the cells, and then centrifugation at 1200RPM for 5mins. Supernatant was discarded and cell pellets stored at -80°C pending purification.

2.4.5 Purification of viruses into high titre stocks

Pellets of cells virus infected cells were thawed and resuspended in 5ml of PBS, 5ml of tetrachloroethylene (TCE, Fisher Scientific) was then added and mixed by vigorous shaking. The mixture was then centrifuged at 2000RPM for 5min in a 15ml falcon tube causing the mixture to separate into layers. The top, aqueous, layer contains the viruses and was collected whilst the supernatant was discarded.

Caesium chloride (CsCl) gradients were prepared in UltraClear 14 x 89mm tubes (Beckman Coulter). 2.5ml of 1.40g/ml CsCl in PBS was added to the tube, 2.5ml of 1.25g/ml CsCl in PBS was slowly pipetted on top to create a separation between the two densities of CsCl. The aqueous phase from the previous TCE purification of the virus pellet was then pipetted, slowly, on top of the CsCl gradient without disrupting the CsCl interface. The remaining volume in the tube was filled with PBS. The tube was placed in an Sw41Ti rotor then centrifuged for 1.5hrs at 18°C at 35600RPM. The tube was carefully removed from the rotor and a series of milky white bands were visible. The uppermost band was cell debris while middle band(s) were immature virions. The lower most band is the mature virus, and was harvested by stabbing a needle through the side of the tube and drawing the band of virus into the attached syringe.

This material was placed on top of a second 5ml CsCl gradient with a single density of 1.34g/ml. the tube was filled with PBS without allowing mixing with the CsCl. This second gradient was centrifuged under the same conditions as above overnight, for a minimum of 15hrs. This time only a single milky white band was visible, which was extracted with a needle in the same manner as previously.

This material was dispensed into dialysis tubing with a 12-14kDa molecular weight cut off (Visking 6.3mm, MediCell Membranes Ltd), and sealed at both ends with Mediclips (MediCell Membranes Ltd). This tubing was placed into 1L of dialysis buffer (10% glycerol, 10mM Tris-HCl (pH7.8), 135mM NaCl and 1mM MgCl₂×6H₂O) at 4°C for approximately 6hrs, the dialysis buffer was then replaced with another 1L and left to dialyse overnight. The dialysed virus containing material was then extracted and divided into 50µl aliquots and stored at -80°C.

2.4.6 Purification of viruses into high titre stocks

Titres of purified virus aliquots were determined based on protein concentration. Total protein concentration was determined using the Pierce Micro BCA protein assay kit (ThermoFisher) according to the manufacturers protocol. In brief, 5µl of purified virus stock was dispensed into a 96 well Nunc plate (ThermoFisher) in triplicate and diluted to 150µl total volume in PBS. 150µl of working reagent was added and incubated in the plate for 2hrs along side the protein standards. The absorbance was measured on an iMark microplate absorbance reader (BioRad) at 570nm.

Data was analysed as follows in order to calculate the viral titre. The blank was subtracted from each of the absorbance readings and then the mean of the three measurements taken. The mean was then multiplied by the dilution factor (30) incurred when the virus stock was diluted up to 150µl of PBS. This value was divided by the coefficient of X for the standard curve (calculated from the protein standards as described in the manufacturers instructions) to give the total protein concentration. The protein concentration in µg/ml was then multiplied by 4x10⁹, the number of viral particles per 1µg of protein. The equation is as follows:

$$\frac{\text{Mean}(\text{Absorbance} - \text{Blank})}{\text{Coefficient of } x \text{ from Standard Curve}} \times \text{Dilution factor} \times 4 \times 10^9$$

$$= \text{Viral particles/ml}$$

This gives the total number of viral particles per ml used as the viral titre throughout this study.

2.5 Infectivity Assays

2.5.1 GFP infectivity assay

Adherent cells were seeded into a Nunc delta surface 96-well cell culture plate (ThermoFisher) at a density of 50×10^3 cells/well in 200 μ l of cell culture media and left to adhere overnight at 37°C in a 5% CO₂ humidified atmosphere. Media was removed and cells washed twice with 200 μ l of PBS. Virus was added to the necessary wells at the required titre in 200 μ l of serum free RPMI 1640 and incubated under cell culture conditions for 3hrs. The virus containing media was then removed and replaced with complete cell culture media and the cells incubated for a further 45hrs. Cell culture media was then removed, the cells washed twice with 200 μ l of PBS, trypsinised in 50 μ l of 0.05% Trypsin-EDTA (Gibco), and dissociated by pipetting. The trypsinised cells were transferred to a 96-well V-bottom plate (ThermoFisher), neutralised with 100 μ l of complete cell culture media, and pelleted by centrifugation at 1200RPM for 3mins. Supernatant was removed, the cells washed once in 200 μ l of PBS, and resuspended in 100 μ l of 2% PFA (PBS containing 2% w/v paraformaldehyde) and incubated at 4°C for 15mins. Cells were again pelleted, washed twice in 200 μ l PBS, then resuspended in 200 μ l PBS prior to analysis by flowcytometry.

Samples were analysed by flow cytometry on Attune NxT (ThermoFisher), voltages were set prior to each experiment, for each cell type, using an uninfected cell population treated identically. Data was 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.

2.5.2 Luciferase infectivity assay

Luciferase infectivity assays were performed using the luciferase assay system kit (Promega). Cells were seeded into a Nunc delta surface 96-well cell culture plate (ThermoFisher) at a density of 20×10^3 cells/well in 200 μ l of cell culture media and left to adhere overnight at 37°C in a 5% CO₂ humidified atmosphere. Media was removed and cells washed once with 200 μ l of PBS. Luciferase transgene encoding replication incompetent viruses were added to the necessary wells at the required titre in 200 μ l of serum free RPMI 1640 and incubated under cell culture conditions for 3hrs. The virus containing media was then removed and replaced with complete cell culture media and the cells incubated for a further 45hrs. Cell culture media was then removed, the cells washed twice with 200 μ l of PBS, and were then lysed in 100 μ l of cell culture lysis buffer (part of the Promega kit) diluted to 1x in ddH₂O. The plate was then frozen at -80°C.

After thaw, 10 μ l of lysate from the cell culture plate mixed then was transferred to a white Nunc 96-microwell plate (ThermoFisher) and 100 μ l of luciferase assay reagent (Promega Kit) added. Luciferase activity was then measured in relative light units (RLU) by plate reader (Clariostar, BMG Labtech). Total protein concentration was determined in the lysate using the Pierce BCA protein assay kit (Thermofisher) according to the manufacturers protocol, absorbance was measured on an iMark microplate absorbance reader (BioRad).

Relative virus infection was determined by normalising the measured luciferase intensity to the total protein concentration (RLU was divided by protein concentration). This gave a final infectivity readout in RLU/mg of protein.

2.5.3 Receptor blocking assay

This assay was also performed using the luciferase assay system kit (Promega). Cells were seeded into a Nunc delta surface 96-well cell culture plate (ThermoFisher) at a density of 20×10^3 cells/well in 200 μ l of cell culture media and left to adhere overnight at 37°C in a 5% CO₂ humidified atmosphere. Media was removed and cells washed 2x with 200 μ l of cold PBS and the plate cooled on ice. 20pg/cell of recombinant adenovirus fiber knob was added to each well in 200 μ l of cold PBS and incubated on ice in a 4°C cold room for 1hr. Media was then removed and luciferase transgene encoding replication

incompetent viruses added to the necessary wells at the required titre in 200µl of cold serum free RPMI 1640 and incubated on ice in a 4°C cold room for 1hr. The virus containing media was then removed and replaced with complete cell culture media and the cells incubated for a further 45hrs under normal cell culture conditions. From this point forward the assay is identical to the one described in section 2.5.2.

2.5.4 Neuraminidase and Heparinase cleavage assays

Cells were seeded at a density of 50×10^3 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. 50µl of neuraminidase (from *Vibrio Cholera*, Merk) at a concentration of 50mU/ml, or 50µl of Heparinase III (from *Flavobacterium heparinum*, Merck) at a concentration of 1U/ml was diluted in serum free media, added to the appropriate wells, 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 5000 viral particles per cell, in 100µl 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 containing PBS for 20mins before wash, and resuspension in 200µl of PBS.

Samples were analysed by flow cytometry on Attune NxT (ThermoFisher), voltages were set prior to each experiment, for each cell type, using an uninfected cell population treated identically. Data was 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.

2.5.5 Competition inhibition assay

CHO cells expressing the appropriate receptor (CAR: CHO-CAR, or CD46: CHO-BC1) were seeded at a density of 30×10^3 cells per well in a flat bottomed 96 well tissue culture plate and incubated at 37°C overnight. Serial dilutions were made up in serum free RPMI 1640 to give a final concentration range of 0.0001-100 $\mu\text{g}/10^5$ cells of recombinant soluble knob protein. Cells were incubated on ice for 30mins, then washed twice with cold PBS. Fiber knob dilutions were then added to the cells and incubated on ice for 30mins. Cells were then washed twice in cold PBS and stained with the primary CAR or CD46 antibody, RmcB (Merk) or MEM-258 (Bio-Rad), respectively, to complex receptors unbound by fiber knob protein, and incubated for 1hr on ice. Cells were washed twice further in PBS and incubated on ice for 1hr with Alexa-647 labelled goat anti-mouse F(ab')₂ (ThermoFisher)^{79,258,260}. All antibodies were used at a concentration of 2 $\mu\text{g}/\text{ml}$.

Samples were run in triplicate by flow cytometry on Attune NxT (ThermoFisher), and analysed using FlowJo v10 (FlowJo, LLC) by gating sequentially on singlets, cell population, and Alexa-647 positive cells. Total fluorescence (TF) was used as the measure of inhibition, rather than percentage of fluorescent cells in the total population, to account for the presence of multiple receptor copies per cell surface which can enable partial inhibition of antibody binding on the cell surface. TF was defined as the percentage of Alexa-647 positive cells in the single cell population for each sample and multiplied by the median fluorescent intensity (MFI) of the Alexa-647 positive single cell population in each sample. Data points are the mean total fluorescence of three biological replicates with error given as the standard deviation from the mean. IC₅₀ curves were fitted by non-linear regression, and used to determine the IC₅₀ concentrations^{79,258,260}. CHO-CAR and CHO-BC1 cells were originally derived by Bergelson *et al*⁶⁹, and Manchester *et al*⁸⁵, respectively.

2.5.6 Statistical analysis of infectivity assays

Figures were generated in GraphPad Prism version 4.03 (GraphPad Software Inc). Unless otherwise states all data is n=3 (See figure legends for specific n numbers), with error given as standard deviation from the mean (SD). P values

are indicated as follows: ns (not significant) indicated $P > 0.05$, * indicates $P < 0.05$, ** indicates $P < 0.01$, *** indicates $P < 0.005$, **** indicates $P < 0.001$.

2.6 Structural Analysis of fiber knob proteins

2.6.1 Homology modelling of HAdV-D8K

The I-TASSER protein structure and function prediction server^{318–320} was used to generate a homology model of HAdV-D8 based on the published sequence of HAdV-D8³²¹, using the published structure of its closest relative (by sequence identity), HAdV-D19p²⁷. The resultant monomer was then copied three times, using the HAdV-19p trimer as a template, and the monomers aligned in PyMol³²² to generate a model of the complete HAdV-D8K trimer.

2.6.2 Crystallisation of fiber knob proteins

Protein samples were purified into crystallisation buffer (10 mM TRIS [pH 8.0] and 30 mM NaCl) during SEC. The final protein concentration was between 5 and 10 mg/ml. Commercial crystallisation screen solutions were dispensed into 3-lens low profile UVP/PS 96-well microplates (SwissSci) using the siting drop format. Drops containing 200nl of protein solution and crystallisation screen at a 1:1, 2:1, and 3:1 ratio were equilibrated against a reservoir of 60 μ l crystallisation solution. The plates were sealed and incubated at 18°C. Crystals appeared between 1hr and 1 week later.

Successful crystallisation conditions for each structure are detailed in the associated data collection statistics tables in chapters 3-5 and in the associated wwPDB depositions.

2.6.3 Soaking with sialic acid

Reservoir solution was made up to a concentration of 30mM N-Acetylneuraminic acid (Neu5Ac, Sigma-Aldrich Cat#A2388) and added to the drop in which the crystals formed in a 1:3 ratio, for a final concentration of 10mM Neu5Ac. HAdV-D26K crystals were incubated overnight prior to harvest, other crystals were soaked in the same manner but for a period of 5mins.

2.6.4 X-Ray structure data collection

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 XDS³²³, Xia2³²⁴, DIALS³²⁵, and Autoproc³²⁶. Scaling and merging data was completed with Pointless³²⁷, Aimless³²⁸ and Truncate³²⁹ from the CCP4 package³³⁰. Structures were solved with PHASER³³¹, COOT³³² was used to correct the sequences and adjust the models, REFMAC5³³³ was used to refine the structures and calculate maps. Graphical representations were prepared with PyMOL³²².

2.6.5 Calculation of electrostatic surface profile

The PDB2PQR server (V 2.1.1)³³⁴ was used to assign charge and radius parameters using the PARSE forcefield, and assigned protonation states using PROPKA, at pH7.35 to approximate the pH of extracellular fluid. APBS³³⁵ was used to calculate electrostatic surface potential, and the map output was visualised in PyMol³²².

2.6.6 Fiber knob to receptor docking

Fiber knob proteins were modelled in complex with CAR or CD46 using the existing HAdV-D37K - CAR liganded (PDB 2J12)⁷³ or the HAdV-B11K - CD46 liganded (PDB 3O8E)⁸⁸ structures, respectively, as a template. Non-protein components and hydrogens were removed from the template model and the fiber knob protein of interest. The two fiber knob proteins were then aligned with respect to their C α chains, in such a way as to achieve the lowest possible RMSD. Models containing only the fiber knob protein of interest and the ligand were saved and subjected to energy minimisation, using the YASARA self-parametrising energy minimisation algorithm as performed by the YASARA energy minimisation server, and results were visualised in PyMol^{322,336}.

2.6.7 Calculation of structure interface energies

Interface energies were calculated using QT-PISA using biological protein assemblies and excluding crystallographic interfaces³³⁷. Values are the mean of the three symmetrical interfaces in each trimer and error is the standard

deviation from the mean, any values above $-3.0 \text{ kcalmol}^{-1}$ were considered to be background as shown as a red dashed line on graphs³³⁸.

2.7 Stable Isotope Labelling of Amino acids in Cells (SILAC)

SKOV-3 cells were cultured in T150 CellBind cell culture flasks (Corning) in DMEM (Dulbeccos modified eagles medium) for SILAC (ThermoFisher) which is deficient in both L-lysine and L-arginine, the media was supplemented with 10% (v/v) of dialysed FBS (ThermoFisher), 100 $\mu\text{g/ml}$ of penicillin-streptomycin (Gibco), L-proline to prevent arginine to proline interconversion, and supplemented with either unmodified L-lysine and L-arginine or Lysine 13C6/15N2 and Arginine 13C6/15N4 in the heavy isotope labelling media. Cells were split in a 1:5 ratio every 72hrs for 2 weeks, for a minimum of 5 passages. Two T150 flasks were seeded with 30×10^5 cells each per condition (i.e. two flasks for the heavy condition and two for the light condition, per sample). The cells were all washed 2x in 10ml of PBS and chilled on ice. PBS was removed from the cells and 50ng of recombinant fiber knob protein, diluted in 5ml of ice-cold PBS, was added to the heavy isotope labelled cells while PBS only was added to the light labelled cells. The flasks were then incubated on a rocking platform at 4°C for 30mins, on ice. The cells were then washed with 10ml of ice-cold PBS, twice, and scraped off the cell growth surface into 5ml of cold PBS and the two flasks per condition combined into a single pellet by centrifugation at 1200RPM for 5min. The supernatant was discarded and 3ml of lysis buffer (0.025M Tris [pH7.4], 0.15M NaCl, 0.001M EDTA, 1% (v/v) NP-40, 5% (v/v) glycerol) added to each sample and incubated with rotation at 4°C for 30min. The samples were then centrifuged at 13000RPM for 20min, the pellet was discarded, and the supernatants of both the heavy and light labelled samples combined into a single tube and 50 μl of Ni-NTA agarose (ThermoFisher) was added. The lysate and beads were incubated together at 4°C for 2hrs then centrifuged at 13000RPM at 4°C for 5min. The agarose pellet was washed three times by adding 1.5ml of lysis buffer, rotating at 4°C for 5min and pelleted by centrifugation at 13000RPM at 4°C. Residual lysis buffer was discarded, and the pellet was resuspended in 30 μl of NuPAGE MOPS SDS running buffer (ThermoFisher) with 0.1M dithiothreitol (DTT) and heated to 100°C for 10min. The sample was pelleted again to separate the Ni-NTA agarose beads. The

supernatant, now containing the pulled down fiber knob protein and any associated proteins it had bound to, was stored at -20°C.

This sample was shipped on dry ice to the commercial vendor FingerPrints Proteomics facility in Dundee, Scotland. This facility performed tandem mass spectrometry (MS/MS) to analyse the peptide content and use the MaxQuant software to determine what proteins peptide sequences are derived from by comparison to a library of human proteins. MaxQuant was then able to calculate a ratio of heavy vs light labelled protein in the sample. For this pull down experiment a 1:1 ratio of heavy to light labelled protein indicates non-specific background interaction between the lysate and the Ni-NTA beads. Enrichment of the heavy labelled proteins indicates that they were associated with the pulldown of the recombinant adenovirus fiber knob protein. The results of the MaxQuant analysis were analysed using the R software environment to determine Altmann Z-scores. Significance was defined as a Z value of >2.0.

2.8 Determination of receptor binding affinity by surface plasmon resonance (SPR)

Binding analysis was performed using a BIAcore 3000™ equipped with a CM5 sensor chip (GE Healthcare). Approximately 5000 RU of CD46, CAR and DSG2 was attached to the CM5 sensor chip, using amine coupling, at a slow flowrate of 10 µL/min to ensure uniform distribution on the chip surface. A blank flow cell was used as negative control surface on flow cell 1. All measurements were performed at 25°C in PBS buffer (Sigma, UK) at a flow rate of 30 µl/min. For equilibrium binding analysis, the HAdV-D26K and HAdV-B3K fiber knob proteins were purified and concentrated to 367 and 3 µM respectively. Five separate 1:3 serial dilutions were prepared for each sample and injected over the relevant sensor chip. The equilibrium binding constant (KD) values were calculated assuming a 1:1 interaction by plotting specific equilibrium-binding responses against protein concentrations followed by non-linear least squares fitting of the Langmuir binding equation. For single cycle kinetic analysis, HAdV-D26K, HAdV-D48K, HAdV-B35K, HAdV-C5K and HAdV-B3K were injected at a top concentration of around 200 µM, followed by four injections using serial 1:3 dilutions. The KD values were calculated assuming Langmuir binding ($AB = B \times AB_{max} / (KD + B)$) and the data were analysed using the kinetic titration

algorithm (BIAevaluation™ 3.1). Receptor proteins were obtained commercially, as follows: Recombinant Human Desmoglein-2 Fc Chimera Protein, R&D systems, Catalogue number 947-DM-100. Recombinant Human CXADR Fc Chimera Protein (CAR), R&D systems, Catalogue number 3336-CX-050. Recombinant Human CD46 Protein (His Tag), Sino Biological, Catalogue number 12239-H08H.

CHAPTER 3. Results: Investigating the broad tropism of species D adenovirus type 49

3.1 Introduction:

The Species D adenoviruses are the largest of the 7 adenovirus species, making up over 70% of their diversity, and are the least well studied^{25,339,340}. Adenovirus vectors which have previously been developed have exploited viral properties which are beneficial to the application at hand. For example, the Enadenotucirev (formerly ColoAd1) vector we selected for its ability to preferentially replicate in colorectal cancer cells, but not in non-transformed cells^{100,291}. The HAdV-D26 vector was selected as a vaccine vector, now in phase III clinical trials for Ebola, due to its strong immunogenicity and subsequent ability to stimulate a lasting immunity^{40,105,341}. Enadenotucirev was selected based on the recombination of various adenovirus species known to efficiently infect colorectal cancer cells, and the immunogenic properties of HAdV-D26 were investigated due to knowledge from epidemiological studies of the species D adenoviruses of the naturally low seroprevalence of wild type HAdV-D26 in the population^{36,40}.

These cases highlight the inescapable truth that regardless of the strategy selected for development of a novel therapeutic vector platform, it must be informed by knowledge of the basic virology underpinning the host range, pathogenicity, and tissue tropism. While cases like the HAdV-D26 vaccine against Ebola¹⁰⁵ and the field of AAV gene therapy³⁴² evidence that complete understanding is not a prerequisite for therapeutic efficacy, both are led by basic understanding of the viruses behaviour in human tissues.

Therefore, in order to investigate the potential of previously understudied species D adenoviruses in vector development we sought to explore the receptor diversity present in a subset of species D adenoviruses. These pseudotypes use HAdV-C5, a well understood adenovirus, as a backbone while expressing the fiber knob proteins of various species D adenoviruses, controlling for differences between adenovirus types mediated interactions apart from the attachment receptor.

During this investigation we observed that the HAdV-C5/D49K pseudotype exhibited the ability to infect cells which were refractory to other adenovirus infection. We further investigated this virus through a series of infectivity assays, and structural comparisons.

HAdV-D49 is a member of the species D adenoviruses for which little is known. It was first isolated from the faeces of a patient with no observable disease, where the prototype T87-677 strain (ATCC VR-1407) was derived from^{343,344}. It has also been isolated from nosocomial epidemic keratoconjunctivitis infections^{345,346}. HAdV-D49 is most associated, however, with patients who are immunocompromised due to HIV infection³⁴⁷. A study of adenovirus infections in patients from the UK and Netherlands found 11 instances of HAdV-D49 infection in 183 HIV positive patients (6% HAdV-D49 positive), compared to just two instances in 2301 tested healthy patients (0.09% HAdV-D49 positive)³⁴⁸.

HAdV-D49 has been of interested previously in the development of therapeutic viruses due to it's very low rates of seroprevalence, thus raising the possibility of circumventing pre-existing immunity in patients. In a cohort of 100 members of the Belgian population only 2% had HAdV-D49 positive sera³⁴⁹. A similar experiment showed none of a 103 patient cohort from Scotland had nAbs against HAdV-D49³⁵⁰, compared to 22% of patients in a cohort of 200 adults from sub-Saharan Africa, highlighting significant geographical variation in seroprevalence³⁷.

Previous studies have suggested that HAdV-D49 may be effective as a vaccine vector. Crucell ltd vectorised HAdV-D49 in order to make it express simian immunodeficiency virus GAG protein (SIVGag). This vector was able to induce strong anti-SIVGag CD8⁺ mediated immunity to SIV, greater than that of a comparable HAdV-C5 vector³⁵¹. Another study sought to exploit this HAdV-D49 vector as a gene therapy to reduce excessive cell proliferation in vascular pathologies due to surgery. This work demonstrated that HAdV-D49 was highly efficient at infecting both endothelial cells and vascular smooth muscle cells, more than the other tested HAdV-C5, D26, B35, or D48 viruses, even after short exposure times³⁵⁰.

Studies in CAR, CD46, and α 2-3 linked sialic acid expressing cells suggested that HAdV-D49 may use CD46 as a receptor, however this seems unlikely based on the receptor studies performed in this thesis (Chapters 3 and 4)³⁵¹. Despite these investigations of HAdV-D49 as a therapeutic vector little is known about its virology, perhaps leading to the lack of recent information about the

vectors progress. This work expands the body of knowledge surrounding HAdV-D49 by investigating its fiber protein (HAdV-D49K) in a controlled context.

3.2 Phylogenetic analysis of the fiber knob domain reveals hidden diversity within the species D

To investigate the degree of diversity present within the major receptor binding determinant, the fiber knob domain, of adenovirus, we performed a phylogenetic analysis. Maximum likelihood tree generation based upon the aligned whole genome sequences of the earliest available strains of each of the canonical adenovirus types separates the adenovirus species into the seven distinct adenovirus species, with numerous sub-clades within the species B and D viruses (Fig3-1A). A separate phylogeny was generated using only the amino acid sequences of the fiber knob domains, derived from the same adenovirus sequences (Fig3-1B). In this tree the relative organisation of the top-level clades containing the seven adenovirus species remained the same. A clear difference is seen, however, in the number of sub-clades within each species. The species B shows less branching when the tree is generated based only on the fiber knob protein (Fig3-1B) compared to the whole genome (Fig3-1A), suggesting the interspecies diversity is dependent on parts of the viral genome other than the fiber knob. The inverse is true of the species D adenoviruses, where we observe a large increase in the number of subclades compared to the whole genome tree. This suggests that not only are the species D adenoviruses the largest species, but they also contain the most diverse set of fiber knob proteins in terms of amino acid sequence.

All the currently available adenovirus fiber knob structures share a similar β -sandwich fold with variable loops, as first described by Xia *et al*⁵³. Wu-Kabat variability analysis confirms that this distribution is conserved across the adenovirus species (Fig3-1C). The plot shows that regions associated with β -strands (by alignment to the known HAdV-C5 structure) have a much lower amino acid variability compared to the intervening regions. This suggests that the diversity in adenovirus fiber knob proteins is primarily within the loops while the core structure is conserved in the β -strands, consistent with previously published structures of fiber knob proteins^{27,53,339}.

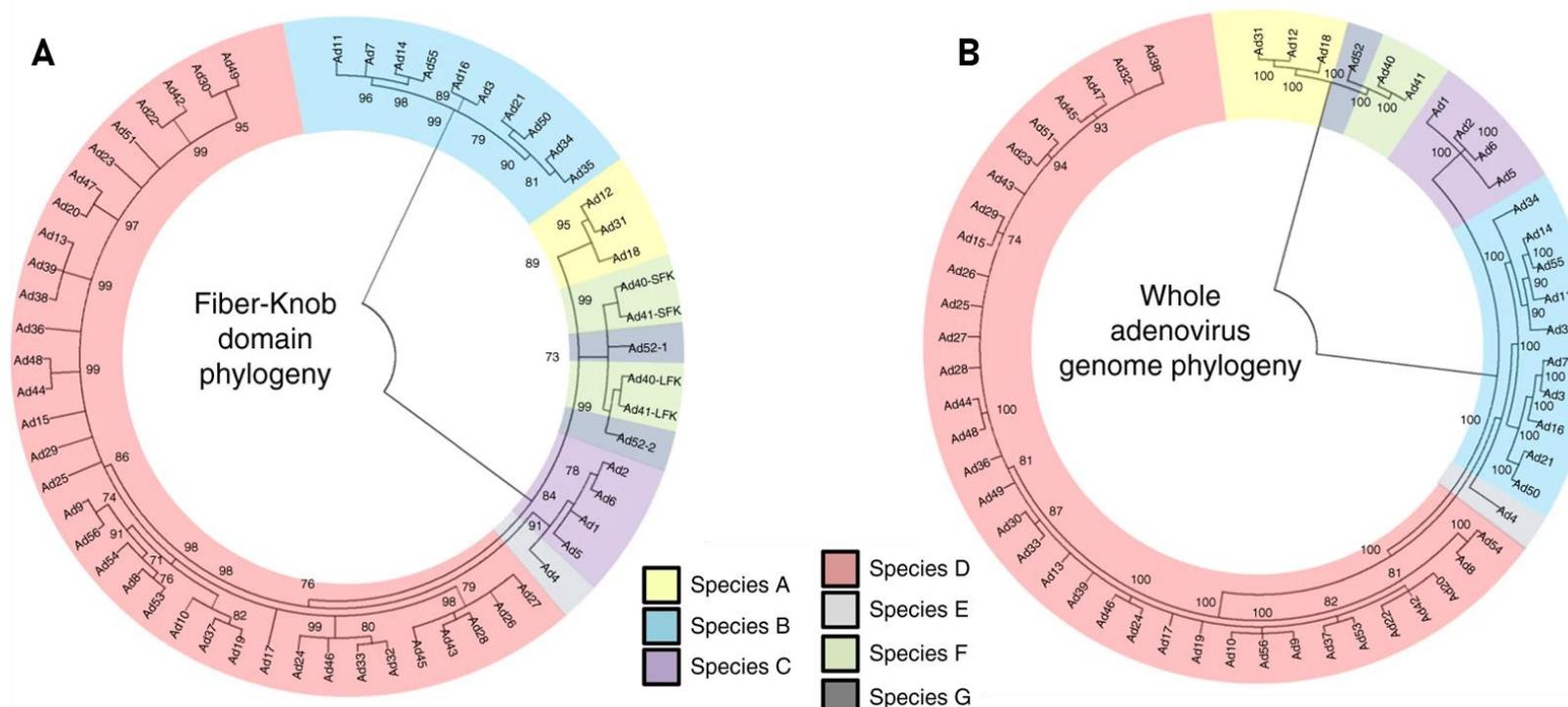
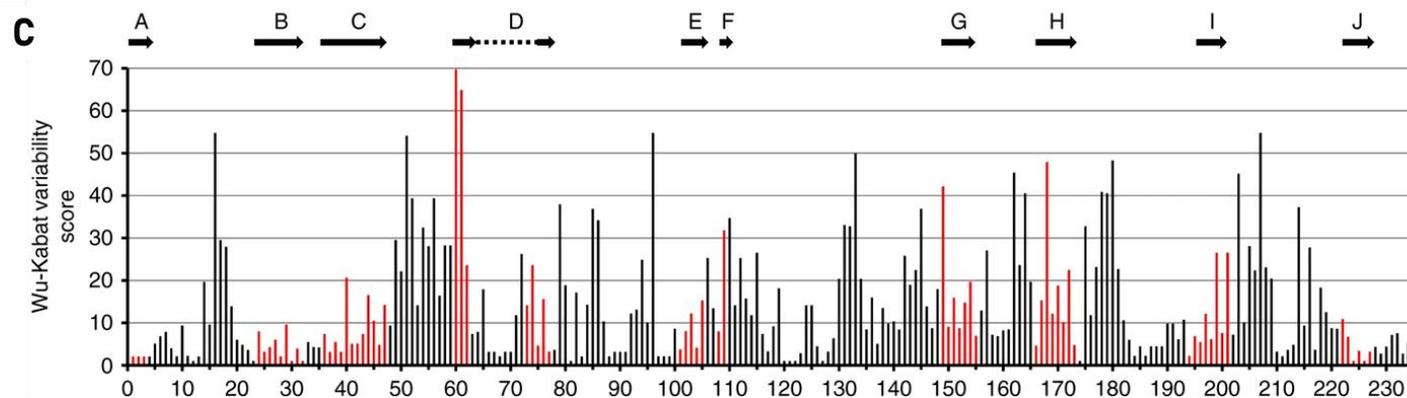


Figure 3-1 Phylogenetic analysis of adenovirus genomes and fiber knob domains reveals increased diversity in the species D adenoviruses.

Maximum likelihood trees generated from the fiber knob sequence of the 56 investigated prototype adenoviruses (A) show different nodes and clade formation compared to trees generated using the aligned whole genomes of these viruses (B). Wu-Kabat variability analysis of the clustal omega aligned fiber knobs of the 56 analysed adenovirus species (C) shows that regions of low variability (lower variability score) correlate with regions known to be involved in β -strand formation in HAdV-C5 (black arrows and red bars).



3.3 Adenovirus fiber knob pseudotypes exhibit different infectivity characteristics in test cell lines

We generated a series of HAdV-C5 based viruses pseudotyped with the fiber knob domain of various other adenoviruses. We investigated the ability of these different adenovirus pseudotypes to infect CHO cells expressing known adenovirus receptors. CHO-K1 cells express no known adenovirus receptors, except for heparan sulfate proteoglycans (HSPGs) which are utilised by HAdV-C5 only when complexed with coagulation factor X^{63,168}, while CHO-CAR and CHO-BC1 cells express human CAR and CD46 (BC1 isoform) respectively. We performed transduction assays utilising replication incompetent HAdV-C5 pseudotypes expressing either Luciferase (Luc) or green fluorescent protein (GFP) as a transgene to investigate virus receptor usage and cell tropism.

We observed that CHO-CAR cells were efficiently transduced by HAdV-C5, which is the prototypical CAR utilising adenovirus, while HAdV-C5/B35K does not infect CHO-CAR cells since HAdV-B35 is known not to interact with CAR in any way (Fig3-2A). We also observed that HAdV-C5/D26K, D48K, D30K, and D49K also infect the CHO-CAR cells with 40-70% efficiency depending upon virus pseudotype, while HAdV-C5/D10K shows no detectable infection (Fig3-2A).

Transduction efficiencies were reversed for HAdV-C5 and HAdV-C5/B35K in CHO-BC1 cells. HAdV-C5 shows 15% transduction efficiency, while HAdV-C5/B35K transduces CHO-BC1 cells very efficiently (approx. 95%) (Fig3-2B). This reflects the known CD46 tropism of the HAdV-B35 fiber knob. HAdV-C5 fiber knob has never been observed to interact with CD46, and the low level of transduction is likely to be background caused by non-specific charge-based interactions between the cell surface and the capsid, direct penton binding to cell surface integrins, or both. HAdV-C5/D26K was shown to achieve around 30% transduction of CHO-BC1 cells, while HAdV-C5/49K was seen to transduce CHO-BC1 cells with similar efficiency to HAdV-C5/B35K. All other tested viruses were seen to infect CHO-BC1 cells at an efficiency similar to, or lower than, the HAdV-C5 negative control (Fig3-2B).

CHO-K1 cells were tested similarly. Commensurate with previous studies³⁵², CHO-K1 cells were shown to be refractory to infection by adenoviruses,

including our pseudotypes, with the exception of HAdV-C5/D49K. HAdV-

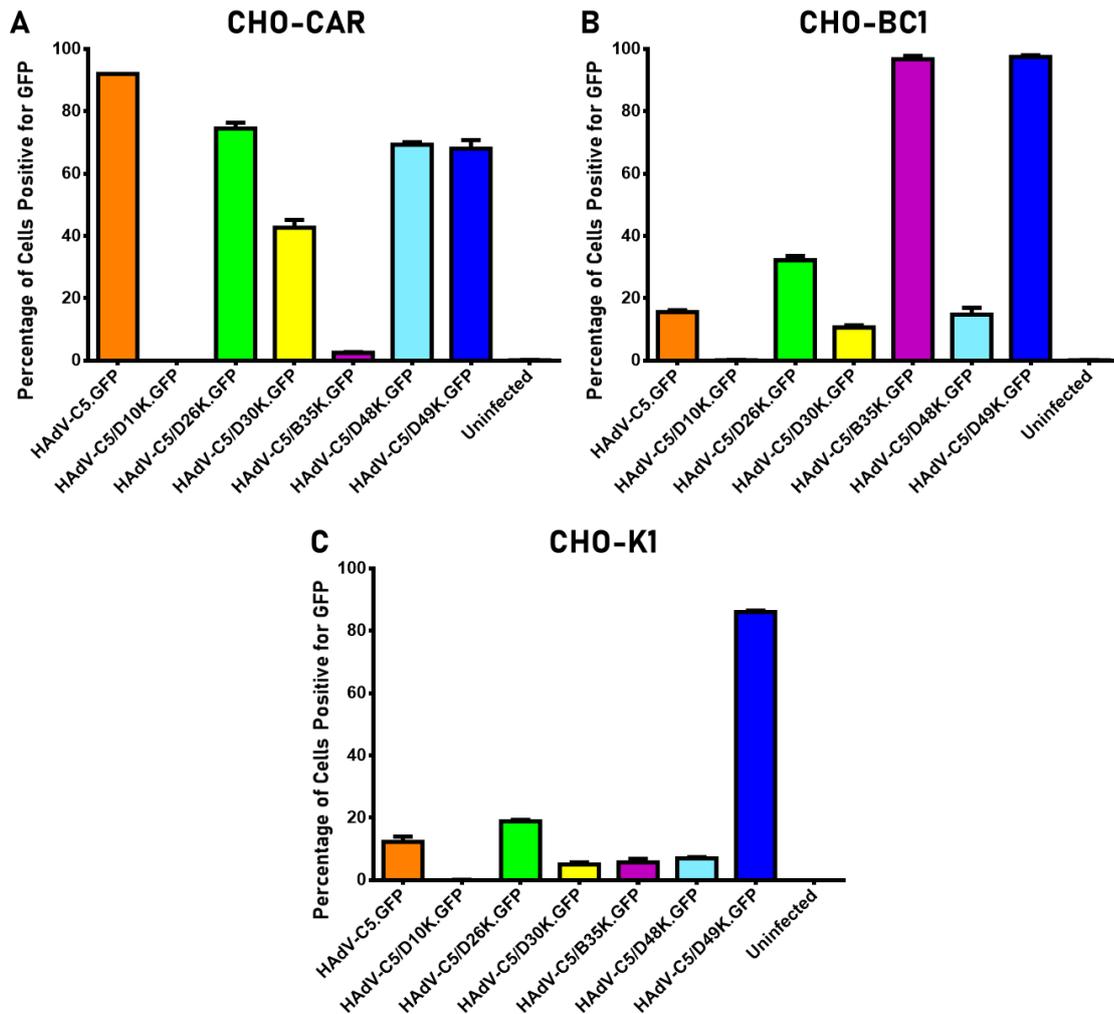


Figure 3-2 Transduction efficiency of adenovirus fiber pseudotypes in CHO cell lines expressing known primary adenovirus receptors, or no known adenovirus receptors. Chinese hamster ovary (CHO) cells expressing human CAR receptor (CHO-CAR, **A**), CD46 receptor (CHO-BC1, **B**), or no known primary adenovirus receptors (CHO-K1, **C**). Cells were infected with 5,000 viral particles per cell with replication deficient HAdV-C5 and the indicated pseudotypes expressing a GFP transgene, $n=3$ error is $\pm SD$.

C5/D49K is capable of transducing CHO-K1 cells, despite them not expressing

any known adenovirus receptor, with an efficiency of approximately 90% (Fig-2C). Interestingly, this virus is less efficient at transforming the CAR expressing CHO cells, than the other types, indicating that CAR may inhibit transduction by the HAdV-D49K virus (Fig3-2).

We wished to rule out the possibility of HAdV-C5/D49K being able to infect cells through an HSPG mediated pathway. As such, we performed a transduction assay in which heparinase was used to remove cell surface HSPGs in CHO-K1 and SKOV-3 cells (Fig3-3). In both cell lines the same pattern was observed. HAdV-C5 was unable to transduce the cells efficiently, due to low levels of CAR expression. However, once the infection media was supplemented with coagulation factor X (FX) HAdV-C5 was able to infect the cells highly efficiently. However, when heparinase has been used to cleave the cell surface HSPGs HAdV-C5 is unable to infect the cells regardless of the presence of FX. By contrast, the transduction efficiency of HAdV-C5/D49K was unaffected by heparinase treatment (Fig3-3A, B), though it is markedly lower than was previously observed in CHO-K1 cells (Fig3-2). This is likely a consequence of the method of the infection assay. In the previous transduction assay (Fig3-2) the virus was incubated with the cells at 37°C for 3hrs, but in order to prevent regeneration of cell surface HSPGs cells were kept on ice and the virus

incubated with the cells for 1hr (Fig3-3). This shorter exposure is likely the cause of the reduced transduction efficiency.

We also needed to rule out whether HAdV-D49K could use another previously established adenovirus cell entry receptor: sialic acid. We performed neuraminidase transduction assays to determine this. As in the heparinase assays, CHO-K1 (Fig3-4A) and SKOV-3 (Fig3-4B) cells were treated with neuraminidase to remove cell surface sialic acid prior to infection. HAdV-C5 transduction efficiency was low, as would be expected due to low CAR expression in CHO-K1 and SKOV-3 cells. The transduction efficiency of HAdV-C5/B35K was low in the non-CD46 expressing CHO-BC1 cells (Fig-4A), but

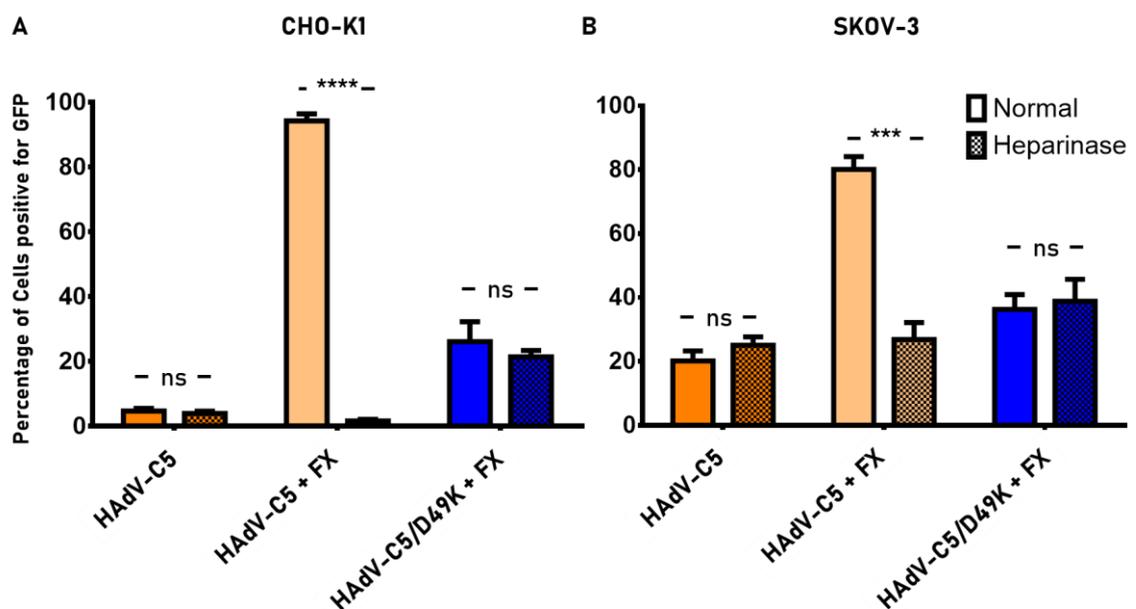


Figure 3-3 Removal of cell surface HSPGs by heparinase treatment does not affect HAdV-C5/D49K transduction efficiency. *The infectivity of HAdV-C5 was unaffected by pre-treatment with heparinase in CHO-K1 (A) or SKOV-3 cells (B) though infection levels were low due to low CAR expression. Infectivity could be restored by the addition of FX during infection and then abolished by heparinase treatment demonstrating dependence upon HSPG expression for efficient HAdV-C5 infection and that heparinase treatment effectively removes cell surface HSPG. HAdV-C5/D49K transduction efficiency was unaffected by heparinase treatment. Cells were infected with 5,000 viral particles per cell with replication deficient HAdV-C5 and the indicated pseudotypes expressing a GFP transgene, n=3 error is \pm SD. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.005$, **** = $P < 0.001$*

high in the CD46 expressing SKOV-3 cells (Fig3-4B), while HAdV-C5/D49K transduction efficiency is the same as observed in the heparinase assay (Fig3-3). When the cells were treated with neuraminidase the HAdV-C5 transduction efficiency was unaffected. The same is true of the HAdV-C5/B35K and D49K pseudotypes in CHO-K1 cells (Fig3-4A). in the SKOV-3 cells HAdV-C5/B35K and D49K transduction efficiency significantly increased (approx. 2-fold) after

neuraminidase treatment for reasons which are unclear. However, it is evident that the removal of cell surface sialic acid does not inhibit HAdV-C5/D49K infection.

To examine the differential infection efficiency in CAR and non-CAR expressing CHO cells, we investigated the ability of the HAdV-D49 fiber knob to affect CAR mediated infection. To do this we utilised a blocking assay in which we used the known CAR utilising HAdV-C5, or the HAdV-C5/D49K to transduce cells which have previously been blocked with either the CAR binding HAdV-C5K, or HAdV-D49K under investigation.

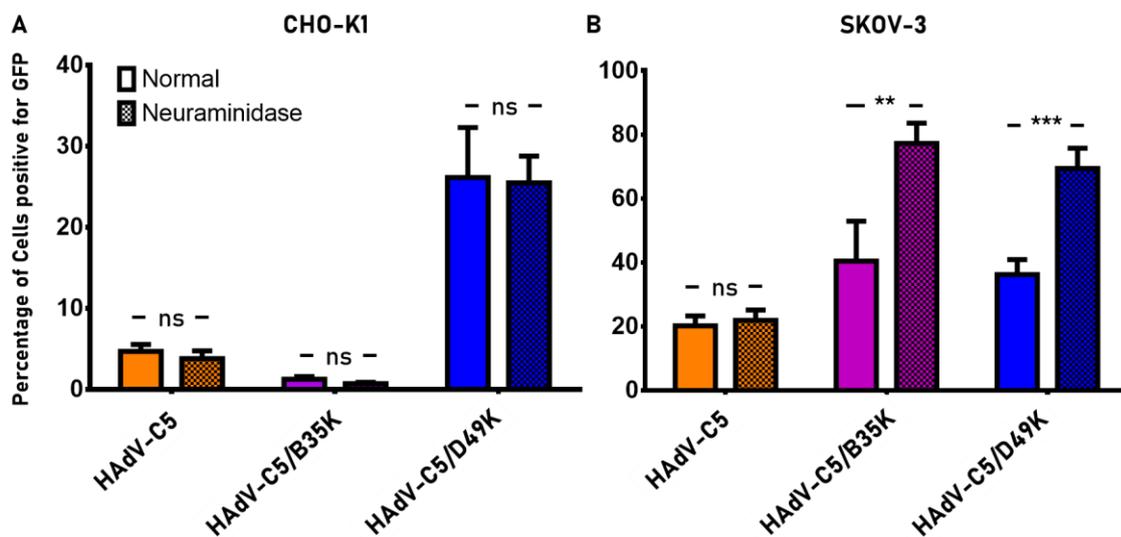


Figure 3-4 Removal of cell surface sialic acid by neuraminidase treatment does not diminish HAdV-C5/D49K transduction efficiency. CHO-K1 (A) and SKOV-3 (B) cells were infected by HAdV-C5, HAdV-C5/B35K, and HAdV-C5/D49K with and without pre-treatment with neuraminidase. None of the viruses were inhibited by neuraminidase treatment, though HAdV-C5/D49K and HAdV-C5/B35K transduction efficiency improved after the removal of cell surface sialic acid for a reason which is unclear. Cells were infected with 5,000 viral particles per cell with replication deficient HAdV-C5 and the indicated pseudotypes expressing a GFP transgene, $n=3$ error is \pm SD. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.005$, **** = $P < 0.001$

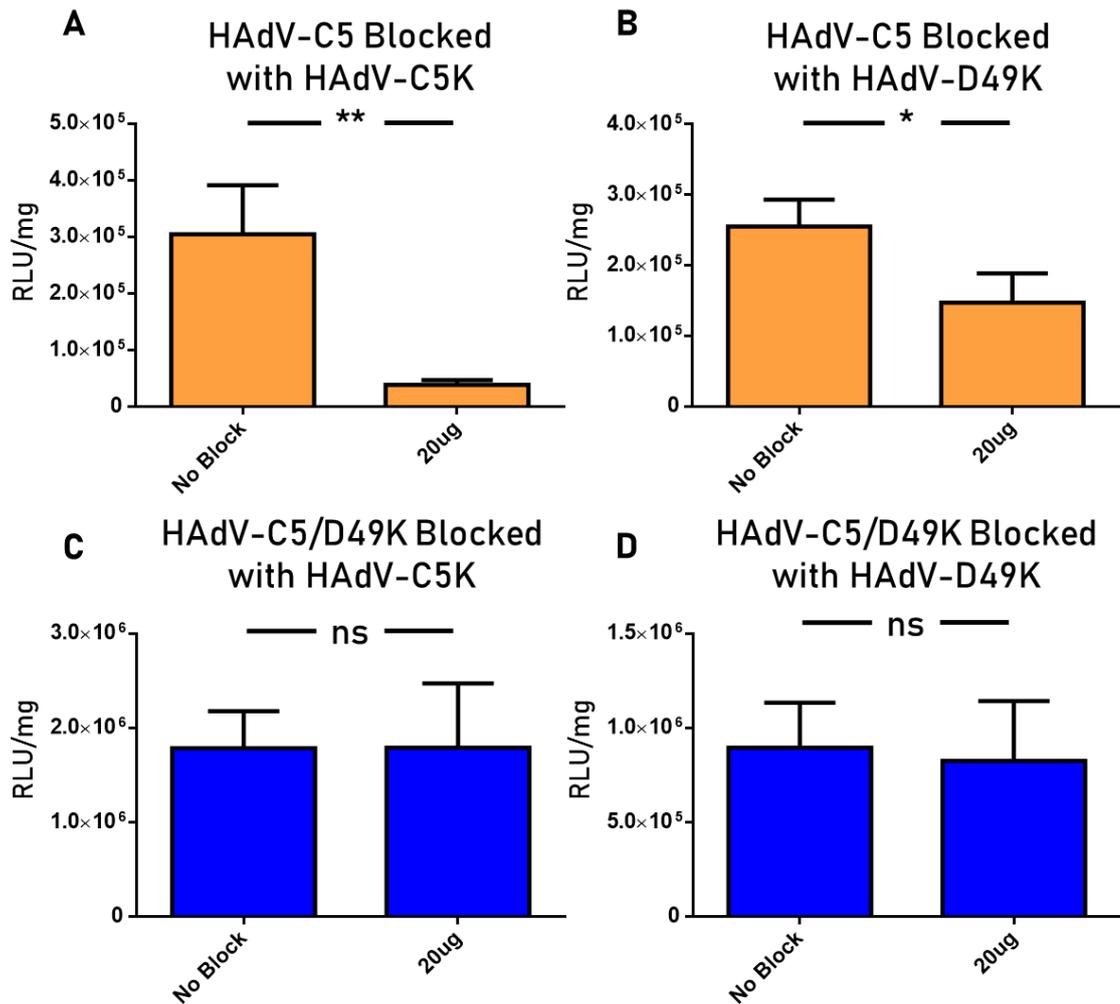


Figure 3-5 HAdV-C5 transduction in CHO-CAR cells is inhibited by HAdV-C5K and HAdV-D49K, but HAdV-C5/D49K is not inhibited by either, suggesting a CAR independent infection pathway. Pharmacological blockage of HAdV-C5 infection by HAdV-C5K significantly inhibited transduction (A), while blockade by HAdV-D49K also inhibited HAdV-C5 infection it was not as efficient (B). The transduction of HAdV-C5/D49K was not significantly inhibited by blockade with either HAdV-C5K (C) or HAdV-D49K (D). Cells were infected with 5,000 viral particles per cell of replication deficient HAdV-C5 or HAdV-C5/D49K expressing a luciferase transgene, with and without blockade by 20 μ g of recombinant HAdV-C5 or HAdV-D49 fiber knob protein. $n=3$ error is \pm SD. * = $P<0.05$, ** = $P<0.01$, *** = $P<0.005$, **** = $P<0.001$

HAdV-C5 infection of CHO-CAR cells was significantly inhibited when blocked

by preincubation with HAdV-C5K (Fig3-5A) or HAdV-D49K, though the latter was less inhibitory (Fig3-5B). However, infection by HAdV-C5/D49K was not significantly inhibited by either HAdV-C5K (Fig3-5C) or HAdV-D49K (Fig3-5D). This highlights a puzzling situation where HAdV-C5/D49K is unable to be blocked by preincubation with its own primary receptor engaging protein.

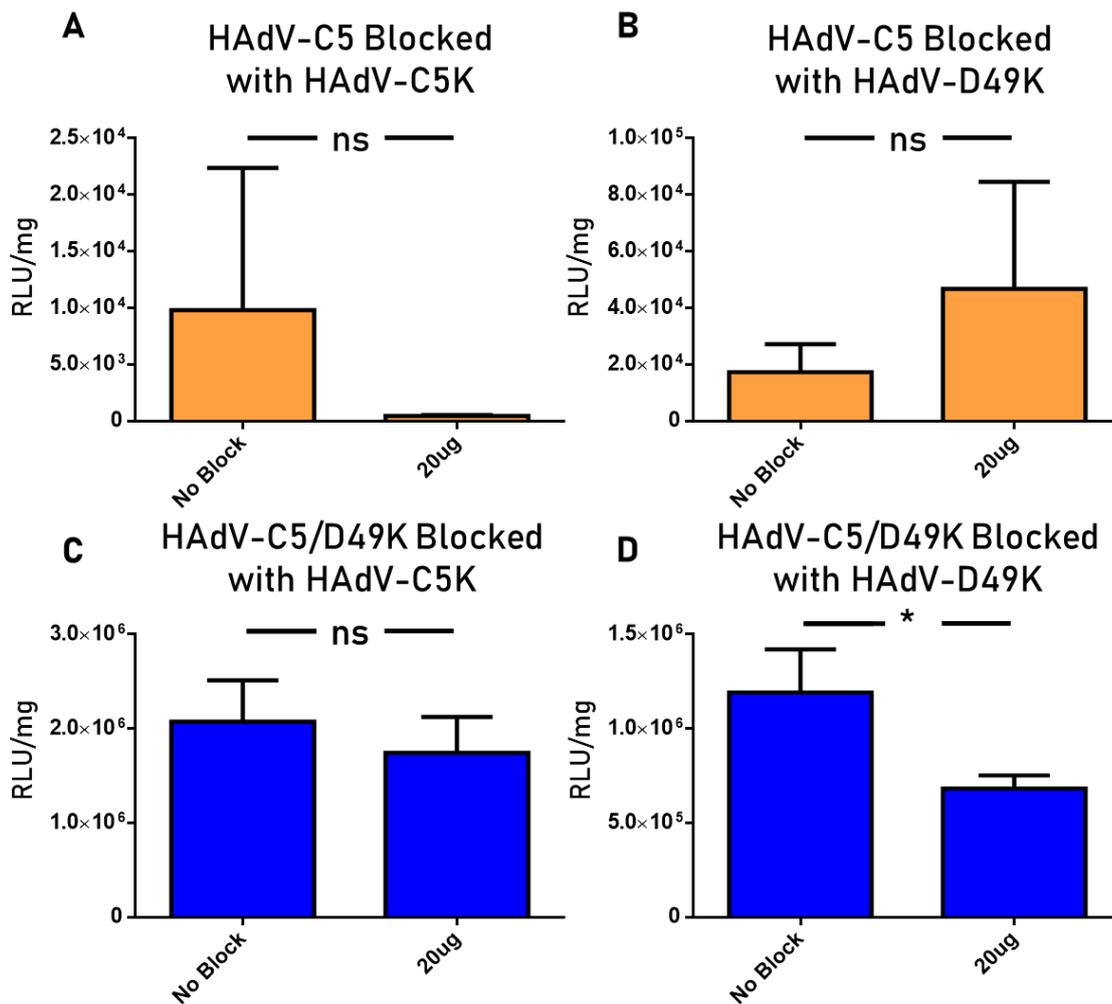


Figure 3-6 HAdV-C5 transduction in CHO-K1 cells is inefficient and not significantly affected by pharmacological blockade, while HAdV-C5/D49K transduction can be inhibited by blockade with its own fiber knob protein. *HAdV-C5 infection was inefficient in CHO-K1 cells and no significant effect was observed when blocked by HAdV-C5K (A) or HAdV-D49K (B). HAdV-C5/D49K transduction efficiently was not affected by HAdV-C5K blockade (C) but was reduced by HAdV-D49K blockade (D). Cells were infected with 5,000 viral particles per cell of replication deficient HAdV-C5 or HAdV-C5/D49K expressing a luciferase transgene, with and without blockade by 20µg of recombinant HAdV-C5 or HAdV-D49 fiber knob protein. n=3 error is ±SD. * = P<0.05, ** = P<0.01, *** = P<0.005, **** = P<0.001*

The same experiments were performed in CHO-K1 cells. No significant differences were observed when HAdV-C5 infection was blocked with either HAdV-C5K (Fig3-6A) or HAdV-D49K (Fig3-6B). Similarly, HAdV-C5/D49K was uninhibited by HAdV-C5K (Fig3-6C). This is unsurprising since very low levels

of transduction were observed due to the absence of CAR receptor. However, HAdV-C5/D49K infection was significantly inhibited in CHO-K1 cells when they were pre-treated with HAdV-D49K, suggesting the ability to interact with a receptor other than CAR on the CHO-K1 cell surface. However, the transduction efficiency of HAdV-C5/49K virus in CHO-K1 cells was still high when blocked with HAdV-C49K, achieving approximately 7×10^5 RLU/mg of protein with blocking compared to 2×10^6 RLU/mg without (Fig3-6D).

3.4 The HAdV-C5/D49K pseudotyped virus infects cancer cell lines with high efficiency

Based on the ability of HAdV-C5/D49 to infect cells which are normally refractory to adenovirus infection we tested whether it would efficiently infect human cancer cell lines, using HAdV-C5 for comparison. We examined a range of nine Pancreatic cancer (Fig3-7), breast cancer (Fig3-8), and lung cancer (Fig3-9) cell lines, as well as an oesophageal, colorectal, and ovarian cancer cell lines (Fig3-9).

In pancreatic cancer cell lines HAdV-C5 transduction efficiency varies widely between approximately 1×10^4 RLU/mg (APSC1 and Panc0403 cells) and 5×10^5 RLU/mg (PT45 cells), while HAdV-C5/D49K varies between 4×10^5 RLU/mg (Panc0403) and 6×10^6 RLU/mg (Panc10). Panc0403 were the least permissive of adenovirus infection for both virus types, though the most similar transduction efficiency when comparing between HAdV-C5 and HAdV-C5/D49K was observed in MIA PaCa2 cells with only a 4.6-fold improvement in transduction by HAdV-C5/D49K. This is in stark contrast to the 210.9-fold enhance transduction efficiency mediated by HAdV-C5/D49K compared to HAdV-C5 observed in BxPc3 cells (Fig3-7).

Consistent with the pancreatic cancer cell lines, breast cancer cell lines were all more efficiently transduced by HAdV-C5/D49K than by HAdV-C5 (Fig3-8). MCF7 cells were the most permissive to HAdV-C5/D49K infection, achieving 6×10^6 RLU/mg, while MDA-MB-231 cells were least efficiently transduced by HAdV-C5/D49K at 1.5×10^5 RLU/mg. MDA-MD-361 and MDA-MD-231 cells only exhibited a 6.7 and 4.6-fold improvement, respectively, in transduction efficiency with HAdV-C5/D49K compared to HAdV-C5. However, the MCF7 and

BT-20 cell lines showed similar improvements in transduction efficiency at 497.3 and 481.2-fold increase, respectively (Fig3-9).

Two tested lung cancer cell lines, NCI-H460 and A549 demonstrated a similar trend observed in the breast and pancreatic cell lines tested. HAdV-C5/D49K demonstrated a similar enhanced transduction efficiency. However, in a third lung cancer line tested, A427 cells, showed HAdV-C5/D49K was less efficient than HAdV-C5 (Fig3-9).

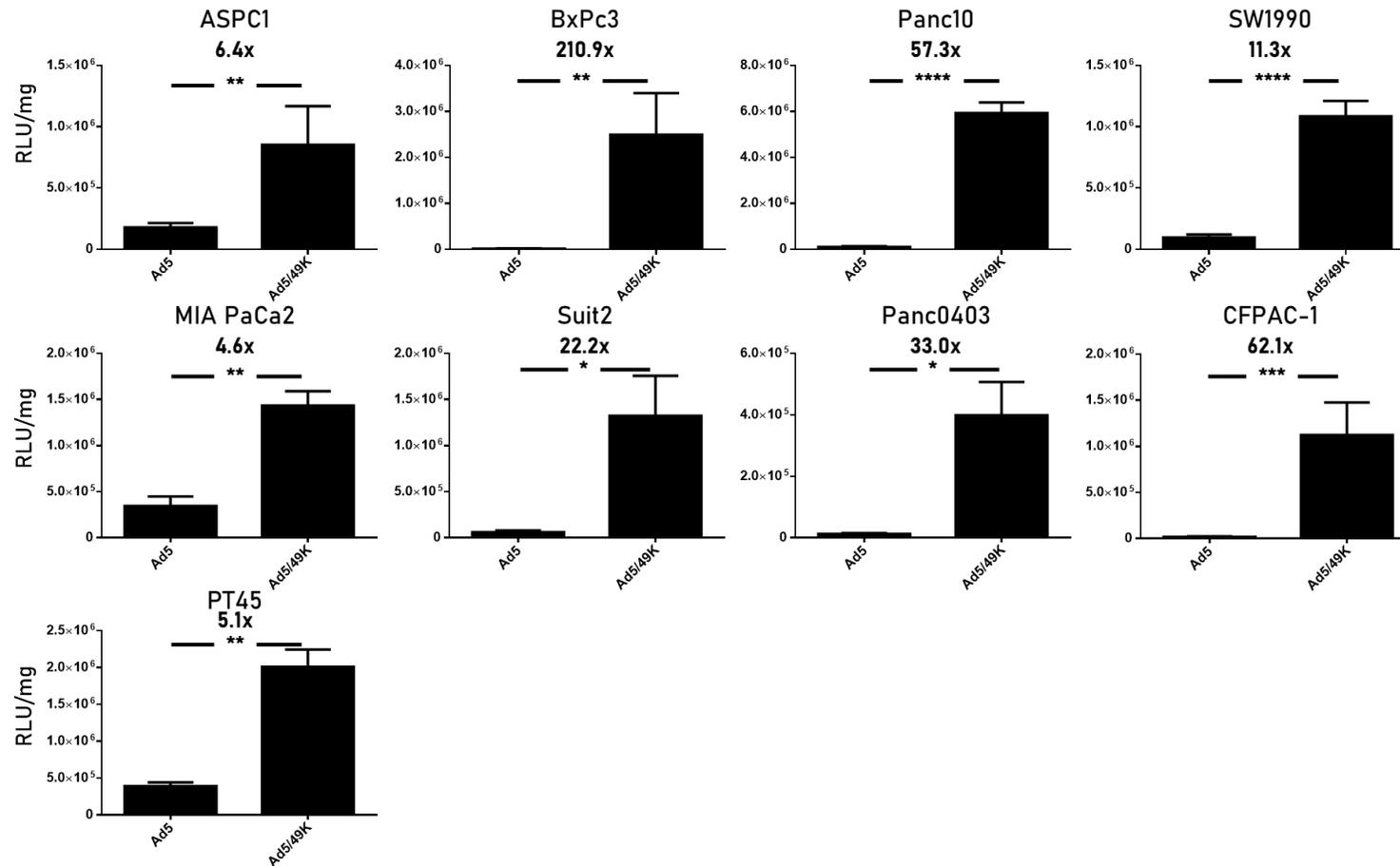


Figure 3-7 HAdV-C5/D49K can transduce human pancreatic cancer cell lines with high efficiency. Transduction was more efficient with HAdV-C5/D49K than with HAdV-C5 in all tested pancreatic cancer cell lines (names given above plots). HAdV-C5/D49K transduction efficiency was measured to be $>10^6$ in all cases, where as HAdV-C5 transduction efficiency varied greatly resulting in variable fold change between the two viruses (boxes below cell line name). Cells were infected with 5,000 viral particles per cell of replication deficient HAdV-C5 or HAdV-C5/49K expressing a luciferase transgene, $n=3$ error is \pm SD. * = $P<0.05$, ** = $P<0.01$, * = $P<0.005$, **** = $P<0.001$**

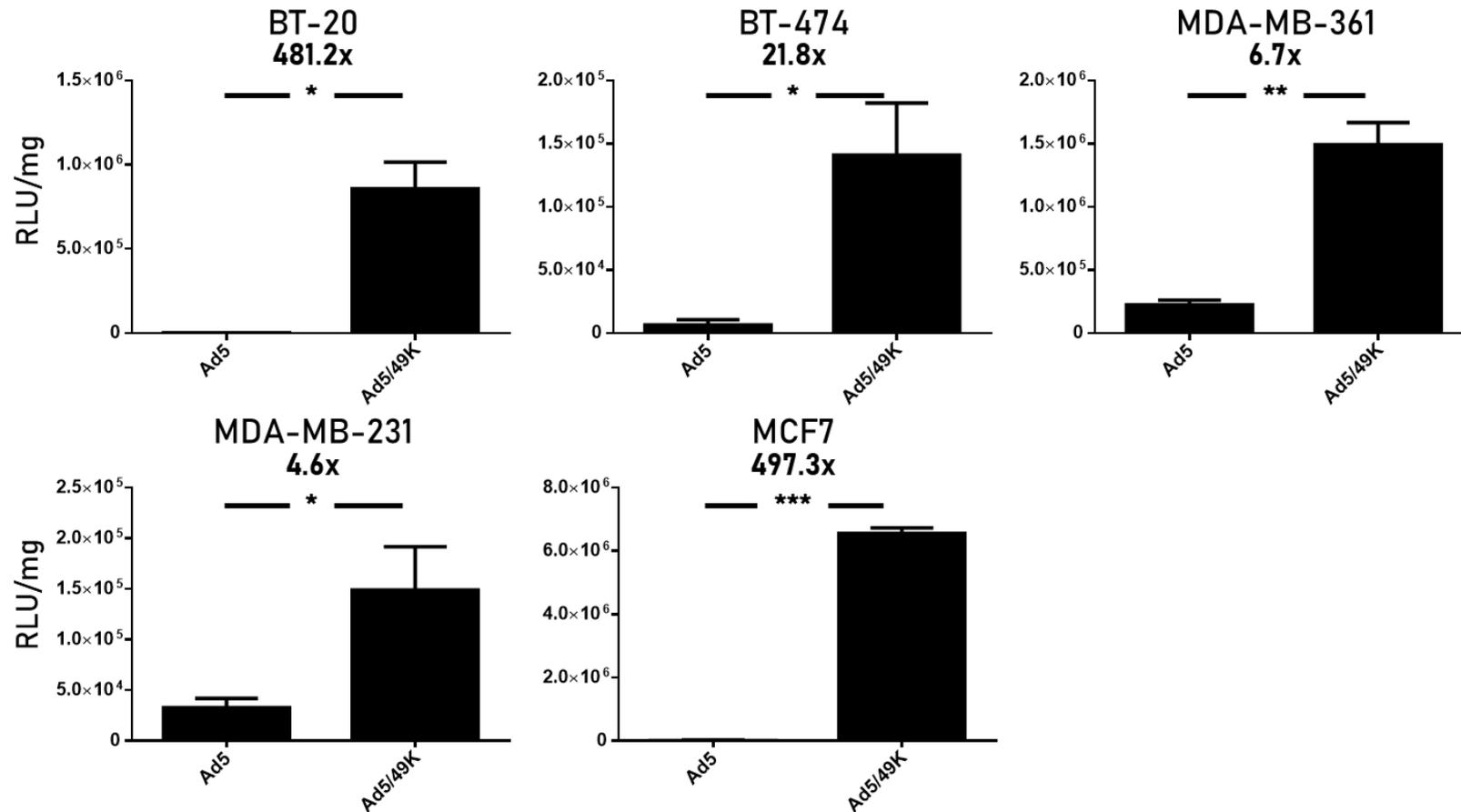


Figure 3-8 HAdV-C5/D49K can transduce human breast cancer cell lines with high efficiency. Transduction was more efficient with HAdV-C5/D49K than with HAdV-C5 in all tested breast cancer cell lines (names given above plots). HAdV-C5/D49K transduction efficiency was measured to be approximately 10⁶, or greater, in all cases, where as HAdV-C5 transduction efficiency varied greatly resulting in variable fold change between the two viruses (boxes below cell line name). Cells were infected with 5,000 viral particles per cell of replication deficient HAdV-C5 or HAdV-C5/49K expressing a luciferase transgene, n=3 error is ±SD. * = P<0.05, ** = P<0.01, * = P<0.005, **** = P<0.001**

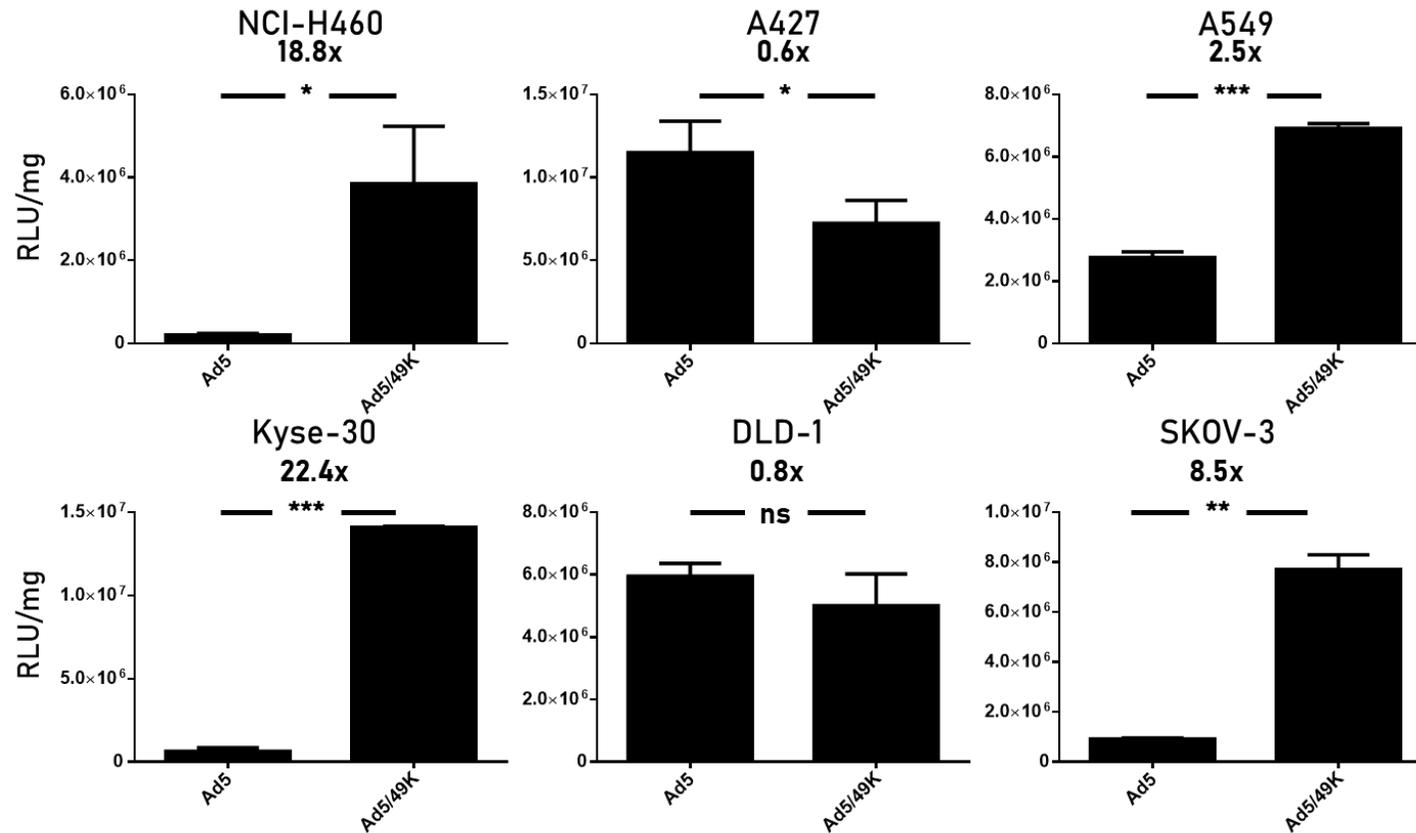


Figure 3-9 HAdV-C5/D49K can transduce human lung, oesophageal, colorectal, and ovarian cancer cell lines with high efficiency. NCI-H460, A427, and A549 cells are lung cancer cell lines, Kyse-30, DLD-1, and SKOV-3 cells are oesophageal, colorectal, and ovarian cancer cell lines, respectively (names given above plots). Transduction was more efficient with HAdV-C5/D49K than with HAdV-C5 in all tested cancer cell lines except for DLD-1 and A427 cells. HAdV-C5/D49K transduction efficiency was measured to be $>5 \times 10^6$ in all cases, whereas HAdV-C5 transduction efficiency varied greatly resulting in variable fold change between the two viruses (boxes below cell line name). A427 cell were transduced slightly more efficiently by HAdV-C5 than by HAdV-C5/D49K, whereas DLD-1 cells transduced with similar efficiency by each virus. Cells were infected with 5,000 viral particles per cell of replication deficient HAdV-C5 or HAdV-C5/49K expressing a luciferase transgene, $n=3$ error is \pm SD. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.005$, **** = $P < 0.001$

The SKOV-3 ovarian cancer cell line and Kyse-30 oesophageal cancer cell lines demonstrated modest improvements in transduction efficiency between HAdV-C5 and HAdV-C5/D49K, compared to the Breast and Pancreatic lines tested (Fig3-9). DLD-1 colorectal cancer cell transformation was more like the A427 lung cancer cells, with HAdV-C5 demonstrating a similar level of transduction efficiency to HAdV-C5/D49K (Fig3-9).

3.5 HAdV-C5/D49K can infect dendritic cells with high efficiency

Previous work has shown adenoviruses to have limited ability to infect dendritic cells (DC)³⁵³. Previous work has also demonstrated that the species D adenovirus HAdV-D26 is an effective vaccine vector (discussed in later chapters)³⁴¹. Given the broad tropism observed for HAdV-C5/D49K we sought to test the vectors ability to transform dendritic cells with the view to investigating its suitability as a vaccine delivery vehicle.

An initial infectivity assay shows that both HAdV-C5 and HAdV-C5/D49K are capable of infecting human monocyte derived dendritic cells (hMDDC, Fig3-10A). HAdV-C5 was observed to transduce 6.7% and 25.2% of the hMDDC at doses of 1000 viral particles per cell at 1K VP/Cell and 5K VP/Cell, respectively. In the same experiment we observed that HAdV-C5/D49K could infect 37.3% and 89.8% of hMDDC at 1K and 5K VP/Cell, respectively. HAdV-C5/D49K therefore represents a 5.6x increase in transduction efficiency of hMDDC at 1K VP/Cell, and a 3.6x increase at 5K VP/Cell, compared to HAdV-C5 (Fig3-10A).

Further investigation of the transduction efficiency reveals larger shifts in the efficiency of transduction than may be expected based upon the percentage of infected hMDDC. We calculated the total fluorescence intensity (TFI), the product of the median fluorescence intensity of the GFP positive hMDDC and the percentage of GFP positive cells, as a relative measure of the total transduction efficiency. Comparing the TFI we observe an 18.1x increase in transduction efficiency between HAdV-C5 and HAdV-C5/D49K at 1K VP/Cell (Fig3-10B). At 5K VP/Cell HAdV-C5/D49K was 16.7x more efficient at transducing hMDDC than HAdV-C5 (Fig3-10C). Therefore, HAdV-C5/D49K is not only transducing more hMDDC than HAdV-C5 is, but it is expressing higher

levels of transgene once it does. Since the transgene expression gene and the rest of the viral capsid are

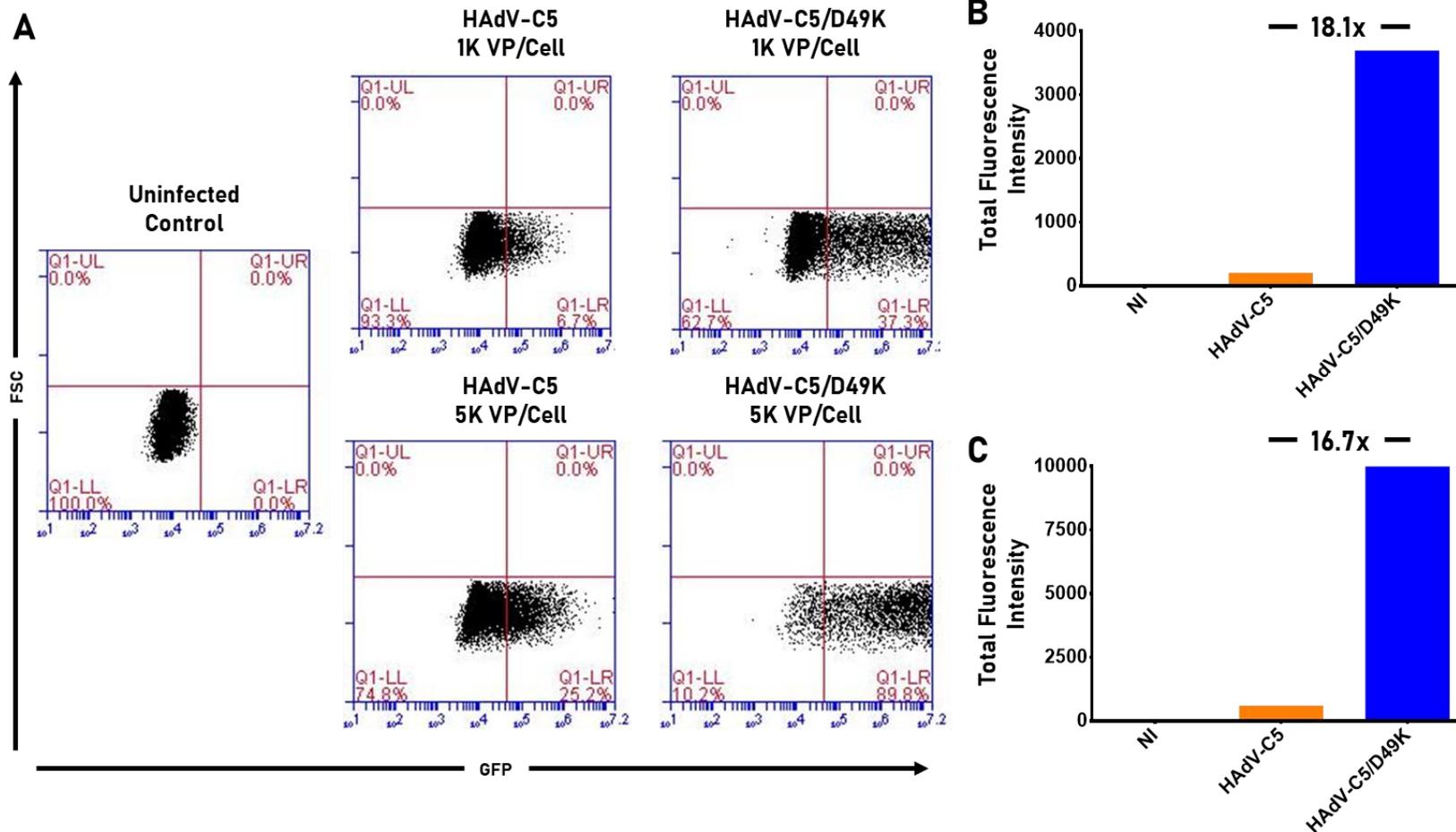


Figure 3-10 Transfection of donor human monocyte derived dendritic cells (hMDDC) is more efficient by HAAdV-C5/D49K than by HAAdV-C5. HAAdV-C5/D49K transduces 30.6% more hMDDC than HAAdV-C5 using 1,000 VP/cell, and 64.6% more when using 5,000 VP/Cell (A). This is equivalent to an 18.1x and 16.7x improvement in transduction when measured in terms of total fluorescent intensity at 1,000 VP/cell (B) and 5,000 VP/cell (C). Cells were infected using replication deficient viruses expressing a GFP transgene.

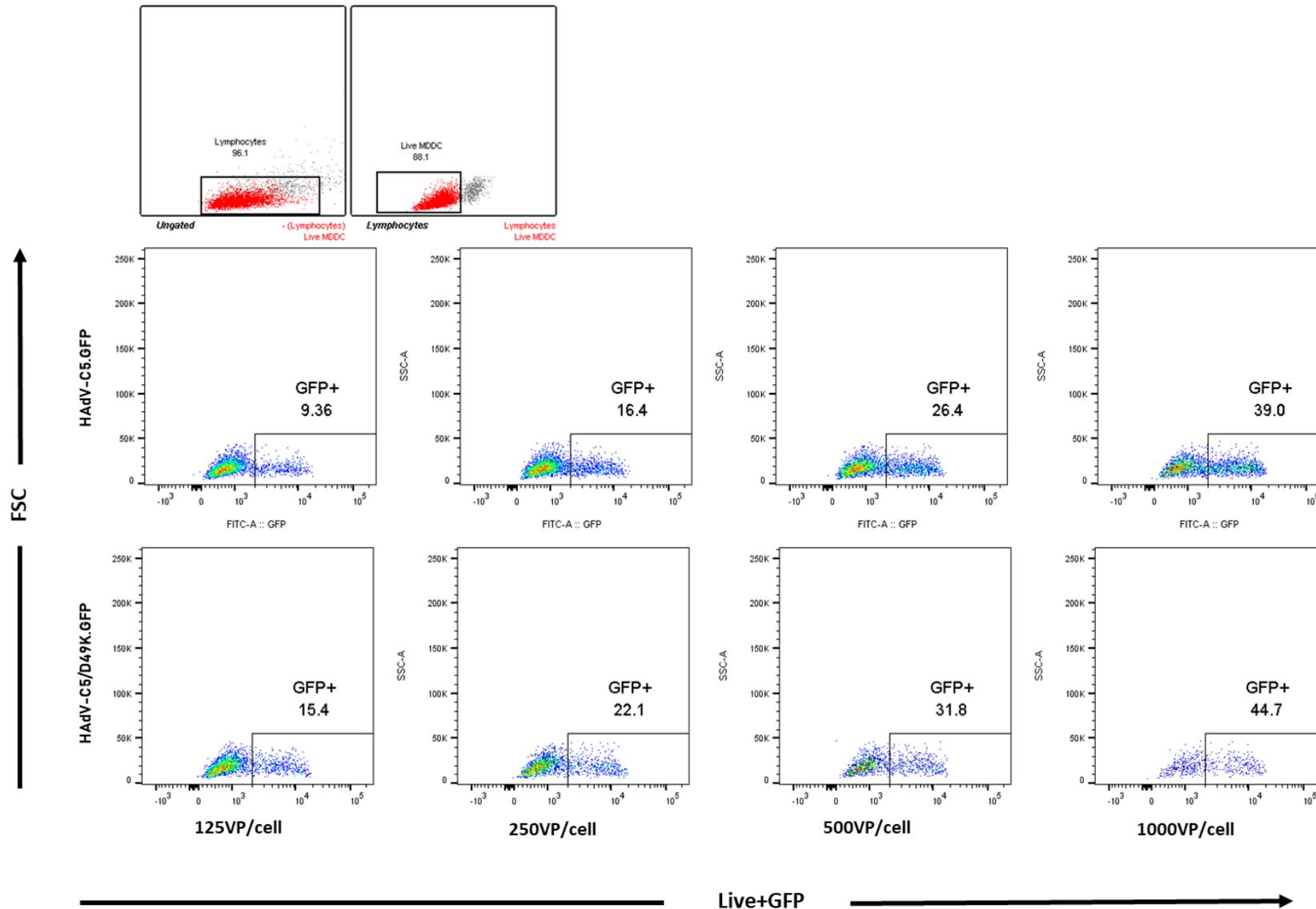


Figure 3-11
 Transduction of donor hMDDC is more efficient using HAdV-C5/D49K, but may result in increased cell death. Transduction of hMDDC was more efficient than with HAdV-C5 at all tested viral doses, however fewer cells are left in the live cell gates for HAdV-C5/D49K than for HAdV-C5 at equivalent doses. Cells were infected using replication deficient viruses expressing a GFP transgene at the indicated doses.

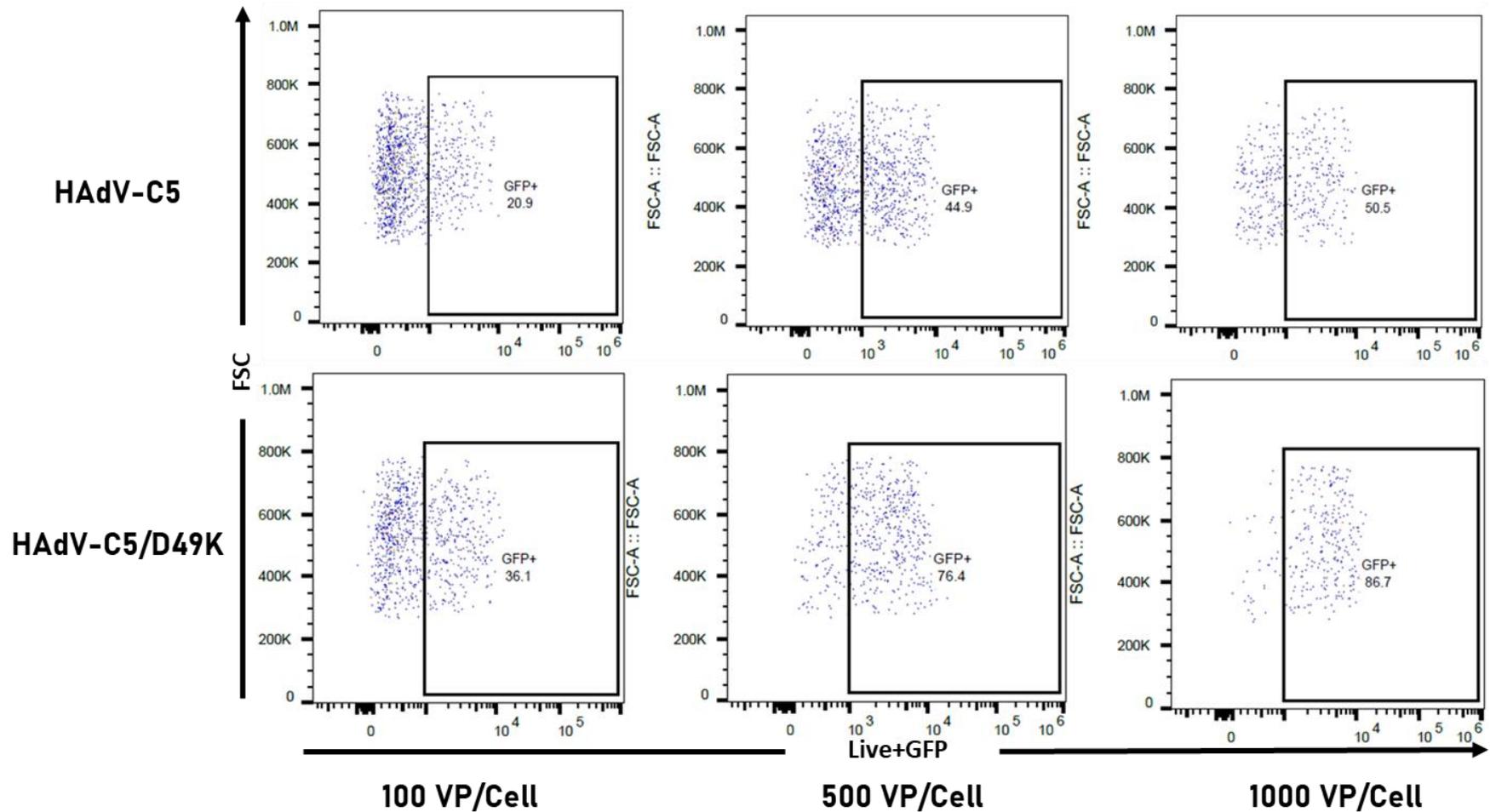


Figure 3-12 Transduction of mouse MDDC shows similar improvement in efficiency using HAAdV-C5/D49K rather than HAAdV-C5. The cells also exhibit greater cell death when infected with HAAdV-C5/D49K than HAAdV-C5 at equivalent doses. Cells were infected using replication deficient viruses expressing a GFP transgene at the indicated doses.

the same between the two viruses and they are replication incompetent the implication is that HAdV-C5/D49K is infecting the GFP positive hMDDC cell population multiple time, delivering multiple transgene copies to yield higher MFI in a given percentage of cells.

A repeat of this experiment, using hMDDC derived from a different donor also showed greater transduction efficiency for HAdV-C5/D49K, but not to the same magnitude (Fig3-11). We observed that increase virus dose results in a larger percentage of cells being transformed, however, higher viral dose is associated with fewer live hMDDC (Fig3-11).

In order to assess the suitability of testing the HAdV-C5/D49K virus in vivo as a DC vaccine vector we sought to test weather this ability to transform hMDDC translated to mouse MDDC (mMDDC). We show a similar trend in the infection of mMDDC as observed in hMDDC, where the HAdV-C5/D49K virus is approximately 40% more efficient than the HAdV-C5 virus at delivering the GFP transgene (Fig3-12).

3.6 The structure of the HAdV-D49 fiber knob domain is highly homologous to that of HAdV-D30 with a different infectivity profile

The data described thus far demonstrates that HAdV-C5/D49K can mediate highly efficient transduction of a range of therapeutically important cell types. CHO cell transduction data suggests that CAR may negatively influence transduction efficiency by the virus (Fig3-2), an effect which was previously observed using whole HAdV-D49 vector³⁵¹, whilst blocking data suggests that it can interact with CAR but is able to infect the cells in a CAR independent manner (Fig3-5,6). However, it remains unclear as to the mechanisms by which HAdV-C5/D49K mediates cellular infection.

Comparison of HAdV-C5 with different pseudotypes shows different transduction efficiencies (Fig3-2). Of interest is the stark difference in transduction efficiency between HAdV-C5/D49K and HAdV-C5/D30K. In CHO-K1 cells HAdV-C5/D49K is extremely efficient at infecting the cells, while D30K shows only background levels of infectivity (Fig3-13A). In CHO-CAR cells both HAdV-C5/D30K and HAdV-C5/D49K infect with approximately 40% and 60%

efficiency, respectively (Fig3-13B). Finally, HAdV-C5/D49K is capable of infecting CHO-BC1 cells with high efficiency (approx. 95%) similar to that of

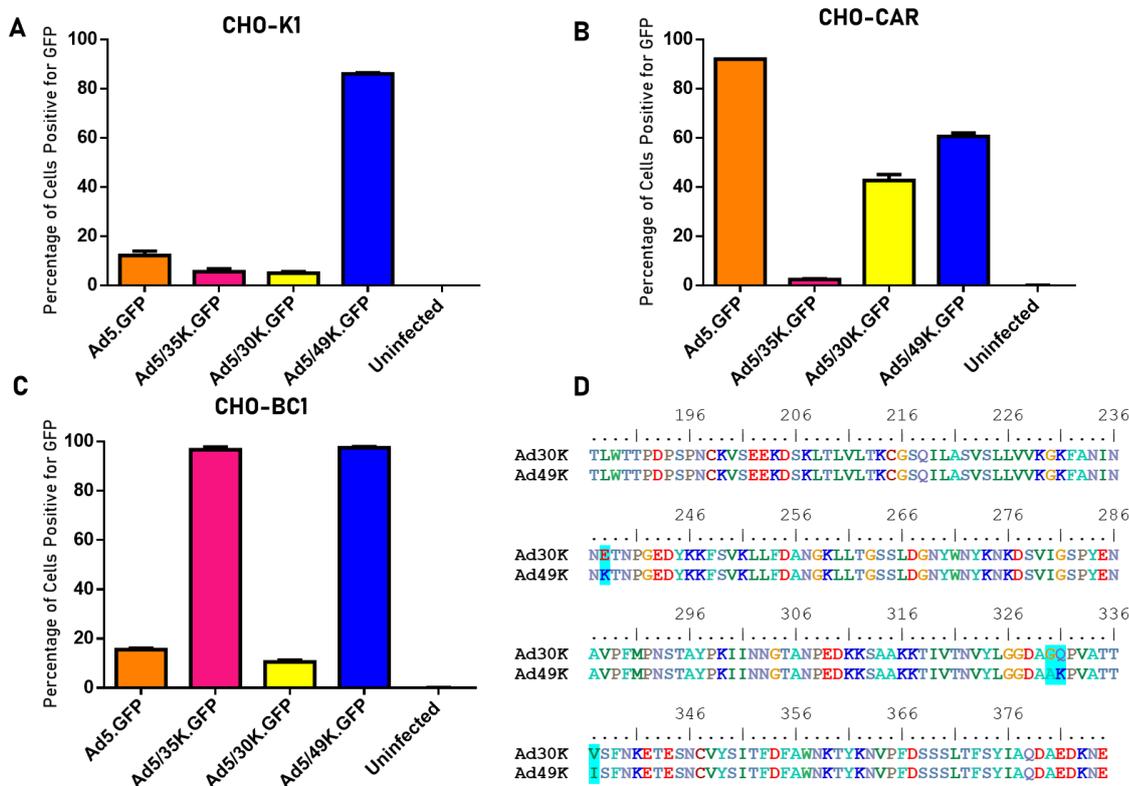


Figure 3-13 HAdV-C5/D49K and HAdV-C5/D30K show different infectivity profiles despite differing in only four residues. HAdV-C5/D49K can efficiently infect CHO-K1 cells while HAdV-C5/D30K is not (A). HAdV-C5/D30K can infect CHO-CAR cells, though not as efficiently as HAdV-C5/D49K or the HAdV-C5 positive control, suggesting the ability to use CAR as a primary receptor with low affinity (B). HAdV-C5/D30K is unable to efficiently transduce CHO-BC1 cells, whilst HAdV-C5/D49K transduces them as efficiently as the HAdV-C5/B35K positive control (C). The HAdV-C5/D30K and HAdV-C5/D49K viruses differ in only four residues, highlighted in blue (D). Cells were infected using replication deficient viruses expressing a GFP transgene at the indicated doses. n=3, error is \pm SD.

HAdV-C5/B35K (Fig3-13C).

Table 3-1 Data collection and refinement statistics for HAdV-D49K (6QPN) and HAdV-D30K (6QPM). One crystal was used for determining each structure. Figures in brackets refer to outer resolution shell, where applicable. MMT (DL-Malic acid, MES monohydrate, Tris: pH 4.0-9.0), SPG (Succinic acid, Phosphate, Glycine). ¹ Coordinate Estimated Standard Uncertainty in (Å), calculated based on maximum likelihood statistics.

PDB Entry	HAdV-D49K – 6QPN	HAdV-D30K – 6STU
Data Collection		
Diamond Beamline	103	103
Date	2017-02-27	18/04/2019
Wavelength	0.97628	0.95372
Crystal Data		
Crystallisation Conditions	0.1M MMT, 25% w/v PEG 1500	0.1M SPG, 25% w/v PEG 1500
pH	8.0	6.0
a,b,c (Å)	106.83, 56.28, 115.70	63.35, 87.36, 217.86
$\alpha=\beta=\gamma$ (°)	90.00, 112.95, 90.00	90, 90, 90
Space group	<i>P</i> 1 2 1	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Resolution (Å)	2.74 – 106.54	2.39 – 54.76
Outer shell	2.74 – 2.81	2.39 – 2.45
R-merge (%)	8.2 (163.3)	0.116 (1.376)
R-meas (%)	9.7 (191.0)	0.134 (1.583)
CC1/2	0.981 (0.400)	0.998 (0.623)
I / σ (I)	6.8 (1.0)	11.2 (1.4)
Completeness (%)	99.1 (98.9)	100.0 (100.0)
Multiplicity	3.7 (3.7)	7.5 (7.8)
Total Measurements	122,219 (9,179)	365,880 (28,110)
Unique Reflections	33,350 (2,461)	48,897 (3,589)
Wilson B-factor(Å ²)	68.9	48.7
Refinement Statistics		
Total number of refined	9,346	9,637
R-work reflections	31,740	46,478
R-free reflections	1,609	2,348
R-work/R-free (%)	22.0 / 26.2	21.7 / 26.8
rms deviations		
Bond lengths (Å)	0.009	0.010
Bond Angles (°)	1.785	1.722
¹ Coordinate error	0.438	0.293
Mean B value (Å ²)	94.2	58.7
Ramachandran Statistics		
Favoured/allowed/Outliers	1121 / 55 / 12	1122 / 78 / 8
%	94.4 / 4.6 / 1.0	92.88 / 6.46 / 0.66

This difference in the ability to transduce cells is surprising as HAdV-D30K is highly homologous to HAdV-D49K, differing in only 4 residues (Fig3-13D). In order to gain greater insight into the properties of HAdV-D49K and the differences with HAdV-30K which grant it this ability to transduce so efficiently we generated crystal structures of both fiber knob proteins (Table 3-1).

HAdV-D30K crystallised with a 2.39Å resolution as 2 biological trimers, 6 monomeric units, per asymmetric unit in a $P 2_1 2_1 2_1$ space group (Fig3-14A). Density was sufficient to clearly define side chain orientations and generate a complete structure (Fig3-14B), and statistics were all acceptable (Fig3-14C). HAdV-D49K also crystallised with a similar arrangement as two biological trimers (Fig3-15A). This structure was in the lower symmetry $P 1 2 1$ space group (table 3-1) suggesting that despite the high sequence homology the packing arrangement is altered by the residue changes observed in figure 13D. The HAdV-D49K data gave similar resolution in the electron density enabling placement of the side chains (Fig3-15B) and the description of a complete structure with acceptable statistics (Fig3-15C).

Inspection of the crystal structures of HAdV-D30K (Fig3-14) and HAdV-D49K (Fig3-15) reveals that the hydrophobic residues (HAdV-D30K - V337 and HAdV-D49K – I337, Fig3-13D) are buried within the β -sandwich domain as part of a β -strand. These residues are both structurally and functionally homologous, stabilising the core fold in a solvent inaccessible region.

Examination of the crystal structures revealed residues which differ between HAdV-D30K (Fig3-16A) and HAdV-D49K (Fig3-16B) are similarly located, as would be expected from the amino acid sequence alignment and high degree of homology (Fig3-13D). The single residue substitution at position 238 is more distant from the 3-fold axis, whereas the dual residue substitution at positions 330 and 331 is closer to the central depression.

Closer inspection of the substituted residues in the context of the HAdV-D30K structure shows that not only are the substituted residues positioned almost identically, but the side chains are similarly oriented (Fig3-16C). This is also observed in the context of the HAdV-D49K structure, which is structurally identical in the region, excepting the 3 residues (Fig3-16D).

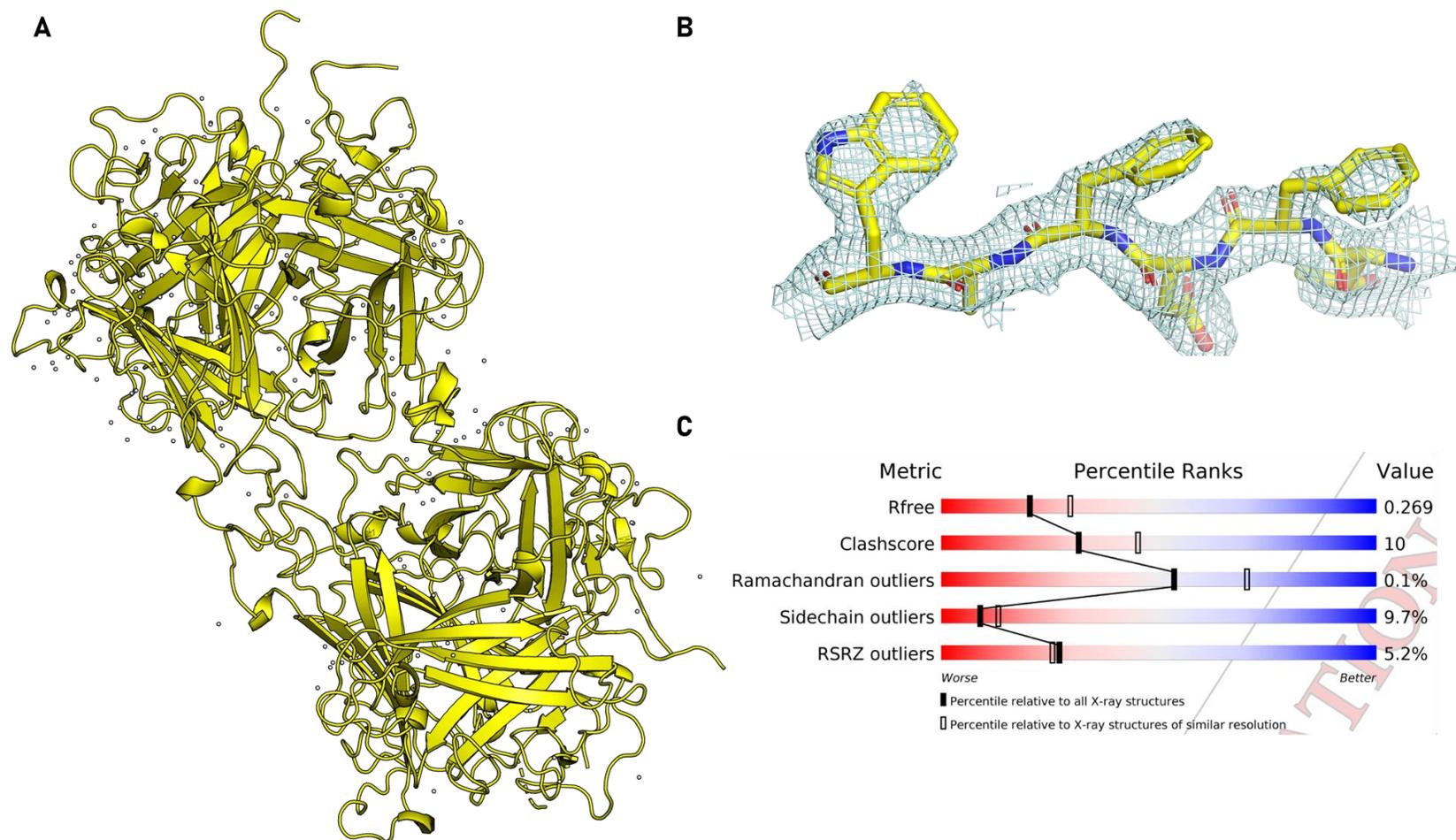


Figure 3-14 Overview of the solved crystal structure of HAdV-D30K, deposited in the wwPDB as 6STU. The asymmetric unit consisted of 6 fiber-knob monomers, creating a pair of trimers, representing two biological units (A), the quality of the map (blue mesh) through the main β -strands was sufficient to define side chain positions (B). The statistics showed scores commensurate with other wwPDB entries at this resolution (C). Map is displayed at $\sigma=1.0$, and carved at 2.0\AA . Protein is shown as a cartoon representation, with missing portions shown as dashed loops.

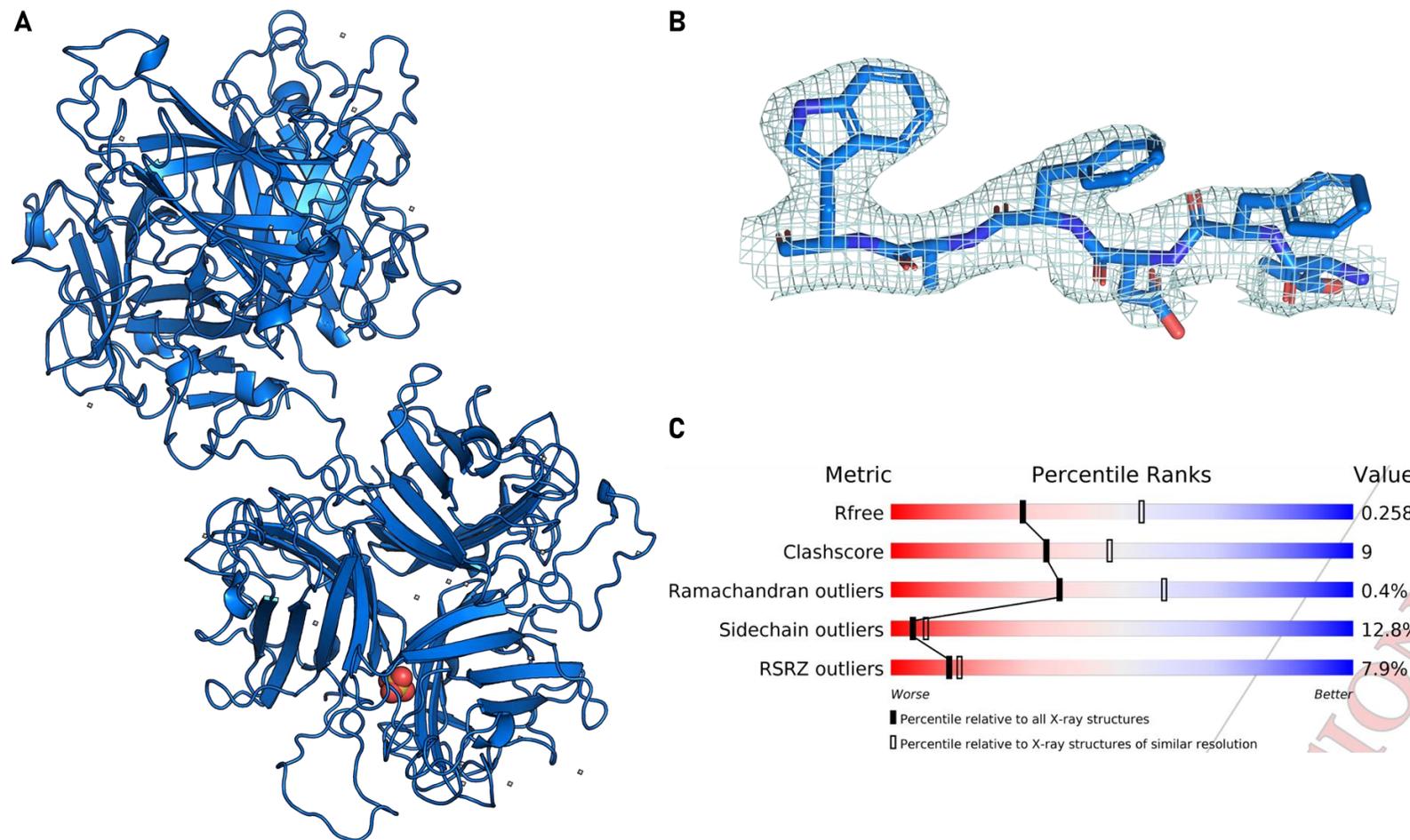


Figure 3-15 Overview of the solved crystal structure of HAdV-D49K, deposited in the wwPDB as 6QPN. The asymmetric unit consisted of 6 fiber-knob monomers, creating a pair of trimers, representing two biological units (A), the quality of the map (blue mesh) through the main β -strands was sufficient to define side chain positions (B). The statistics showed scores commensurate with other wwPDB entries at this resolution (C). Map is displayed at $\sigma=1.0$, and carved at 2.0\AA . Protein is shown as a cartoon representation, with missing portions shown as dashed loops.

In both crystal structures the different residues at amino acid positions 238, 230, and 231, are at the apex of two loops (Fig3-16C, D). The amino acid side chains are oriented to be solvent facing. The swap from E238 in HAdV-D30K to K238 in HAdV-D49K swaps a long δ^- glutamate residue for an even longer positively charged lysine residue. The exchange of glutamine at position 331 in HAdV-D30K trades the polar, but overall charge neutral, residue for a positively charged lysine in HAdV-D49K. The final glycine 330 in HAdV-D30K is swapped for alanine in HAdV-D49K, adding a small hydrophobic side chain to HAdV-D49K. Overall this adds positive charge at the apex of the fiber knob of HAdV-

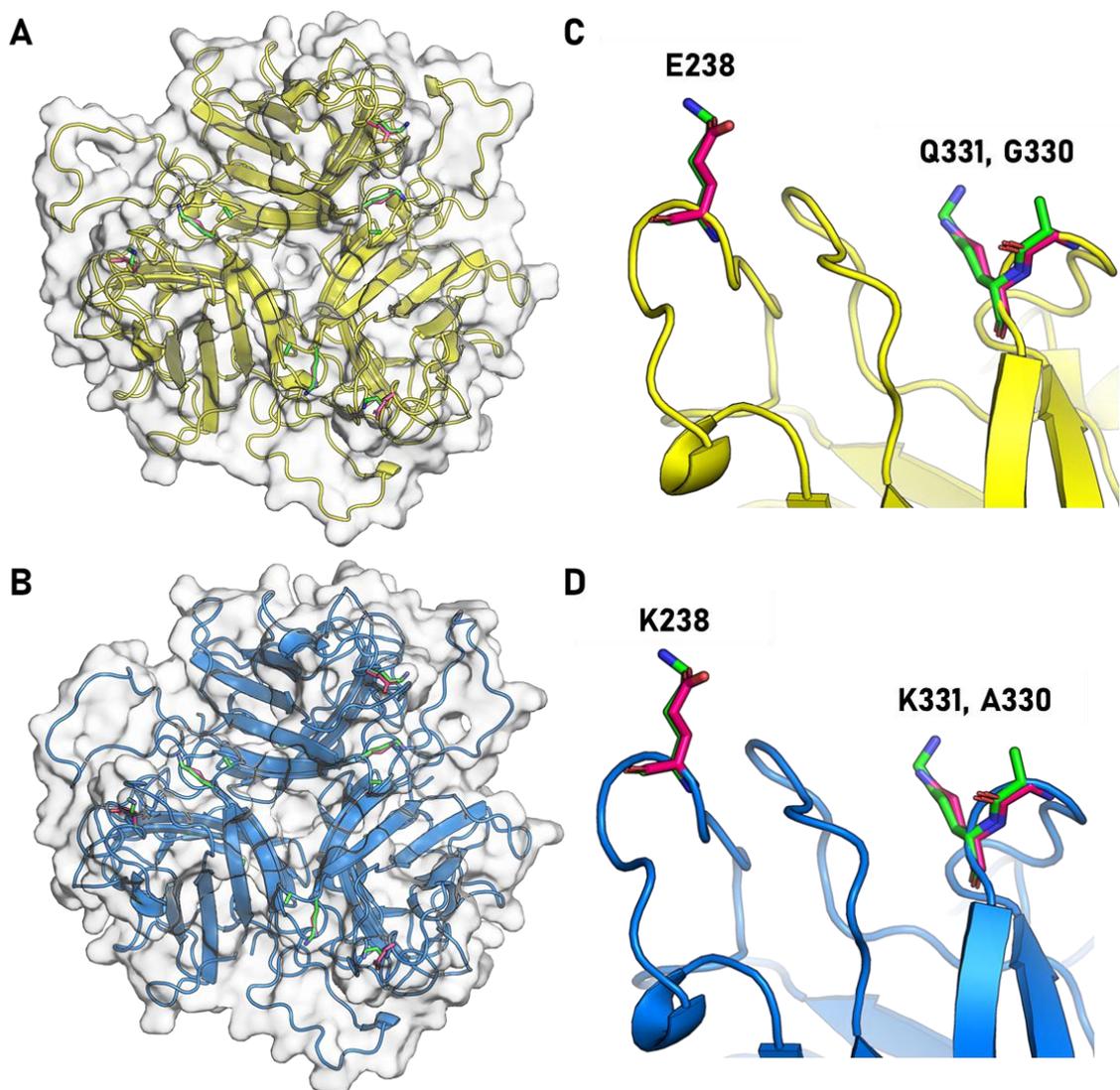


Figure 3-16 Non-conserved residues in HAdV-D30K and HAdV-D49K are similarly oriented in space at the apex of the fiber knob. Viewing the fiber knob protein from the apex, down the 3-fold axis, as it would project away from the virus capsid, we see the residues which differ between HAdV-D30K (pink sticks, **A**) and HAdV-D49K (green sticks, **B**). These residues can be seen projecting into the solvent from loops on the apex of HAdV-D30K (yellow, **C**) and HAdV-D49K (blue, **D**), residue numbers and names match those as they occur on the fiber-knob protein depicted in that frame.

D49.

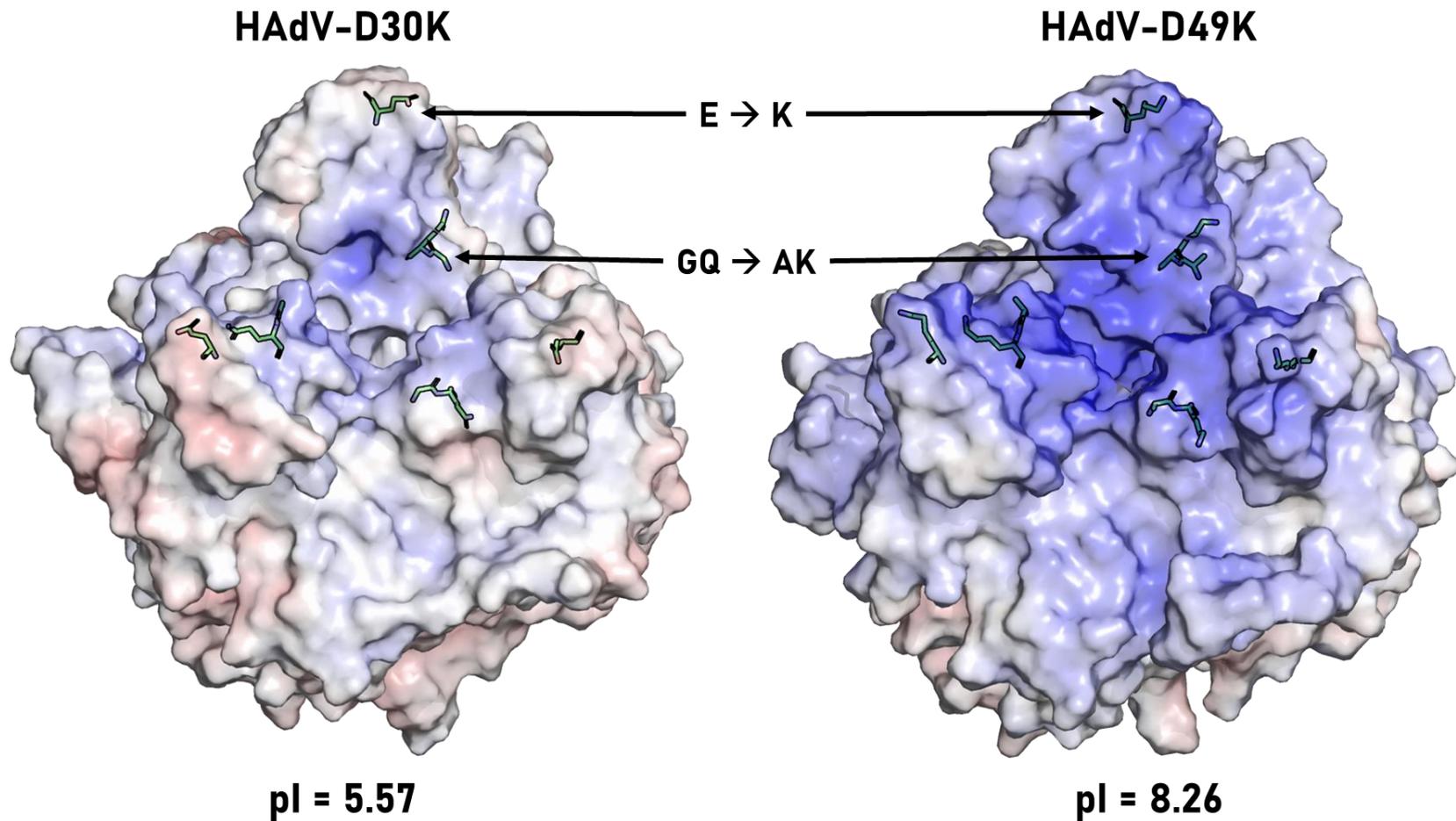


Figure 3-17 The differences in residues between HAdV-D30K and HAdV-D49K have a large effect on the surface electrostatic potential of the fiber knob. The calculated pI of HAdV-D30K and HAdV-D49K is very different as a result of the residue changes, which are shown as green sticks as they occur in that fiber knob protein. The calculated electrostatic surface potential at pH7.35 is projected on a -10mV to +10mV ramp (Red to Blue). HAdV-D49K can be seen to have much more basic potential around the apex where the residue substitutions are.

To investigate the effect of these residue changes on the surface charge we calculated the electrostatic surface potential of both HAdV-D30K and HAdV-D49K (Fig3-17). Though the sequence is only different in 3 surface exposed amino acids there is a large shift in the isoelectric point (pI). HAdV-D30K has a predicted pI of pH5.57 compared to pH8.26 for HAdV-D49K, a nearly 3-log shift, making HAdV-D30K a net acidic protein and HAdV-D49K net basic. Since the difference in charge is due to the difference in bases, it makes sense that the difference in electrostatic potential would be concentrated at the apex of the protein. When the electrostatic surface potential is projected onto the protein surface (Fig3-17) this is what we observe. HAdV-D30K is slightly basic within the central apical depression but is surrounded by a “crown” of negative charge caused by the glutamate residue at position 238 on the surrounding loops (Fig3-17C,D). By contrast HAdV-D49K has a much stronger basic charge within the central depression, and this positive surface potential extends out to the crown ending at lysine 238, the most apical residues on the protein.

3.7 SILAC Co-precipitation suggests that HAdV-D49K may interact with an adhesion protein

To attempt to determine the primary receptor partner is for HAdV-D49K, we performed stable isotope labelling of cells (SILAC) and co-precipitation (Co-precipitation). In this study we labelled SKOV-3 cells by growing them in the presence of amino acids made of either heavy or light carbon isotopes. We then incubated the heavy labelled cells with recombinant HAdV-D49K fiber knob protein. The heavy and light cells were then lysed and combined together, then the HAdV-D49K and any binding partners precipitated onto Ni-NTA agarose beads to extract the fiber knob and its binding partner. The extract was then examined by mass spectroscopy, where we can determine the heavy to light ratio of peptides belonging to different proteins. In theory, peptides with an equal ratio of heavy and light peptides should be non-specific background interactions, and specific HAdV-D49K interactions should be enriched by interaction and pulldown with the fiber knob protein.

We observed that peptides belonging to five unique proteins were enriched in

the SILAC Co-precipitation experiment (Fig3-18). The lowest significance of

these, with a Z-score of 2.10 (Cut off for significance was $Z=2.0$), was dual

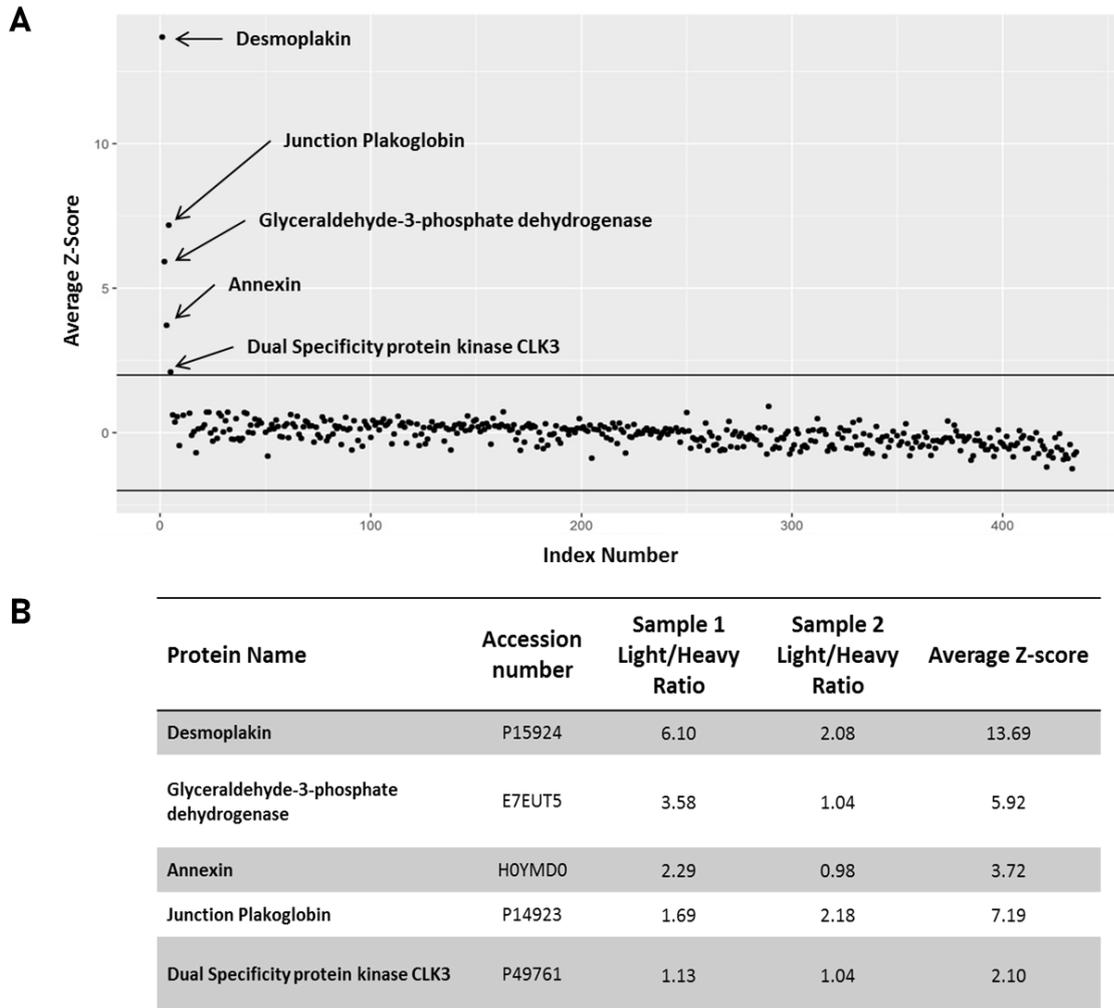


Figure 3-18 Proteins enriched upon affinity pull down assay performed with recombinant rAd49kn protein as determined by SILAC mass spectrometry in labelled SKOV-3 cells. Scatter plot shows the average Z-scores of biological replicates vs the protein index number, assigned arbitrarily (**A**). Solid line represents $Z=2.0$. The table shows the 5 significant hits which were significantly enriched (Mean $Z>2.0$) across both biological replicates of the rAd49kn SKOV-3 pull down assay and the associated protein identifying accession number, as well as the Z-score in each sample (**B**).

specificity kinase CLK3 (CLK3, Fig3-18B). CLK3 a kinase protein involved in

alternative mRNA splicing. Annexin, a group of proteins involved in cytoskeletal organisation and vesicle trafficking, was also found to be moderately enriched ($Z=3.72$) in the SILAC Co-precipitation experiment. With a Z-score of 5.92 Glyceraldehyde 3-phosphate-dehydrogenase (GAPDH) was the third most enriched protein and is involved in glycolysis, transcription regulation, and ER-Golgi transport (Fig3-18B).

The two most enriched proteins were Desmoplakin (DSP, $Z=13.69$) and Junction Plakoglobin (JUP, $Z=7.19$). Both of these proteins are involved in the formation of desmosomes, a form of intercellular junction key to the normal function of cardiac and epithelial tissues³⁵⁴. None of the pulled down proteins are cell surface localised, so it seems implausible that they are the primary cell surface receptor for HAdV-D49K. However, these proteins are key to processes involved in viral trafficking and in the formation of cell-cell junctions, which are known targets of adenovirus infection as both CAR and DSG2 are involved in tight junction formation.

3.8 Discussion

This study explores the natural diversity within the species D adenovirus fiber knob proteins and seeks to understand their impact on infectivity. We first explored the degree of phylogenetic diversity in adenovirus fiber knob domains (Fig3-1), then used pseudotyped HAdV-C5 virus expressing different species D fiber knob proteins to investigate the capacity of those proteins to engage known adenovirus receptors, CAR and CD46 (Fig3-2A,B). One of the virus pseudotypes, HAdV-C5/D49K (Adenovirus type C5 expressing the fiber knob protein of adenovirus type D49), displayed the surprising ability to infect cells which do not express any known adenovirus receptors (Fig3-2C). Further investigation of HAdV-C5/D49K revealed that the virus was able to interact with CAR, as evidenced by the ability of the HAdV-D49 fiber knob protein to inhibit infection by the CAR utilising HAdV-C5 (Fig3-5), it was not dependent upon CAR for infection (Fig-4). Based on this broad ability to infect cells it was compared to HAdV-C5, which is widely used in clinical trials, in pancreatic (Fig3-7), breast (Fig3-8), lung, oesophageal, colorectal, and ovarian cancer cell lines (Fig3-9). Here, HAdV-5/D49K was shown to infect the various cancer lines more efficiently than HAdV-C5 in 18 of the 20 lines tested. HAdV-C5/D49K was also found to infect both human (Fig3-10) and mouse (Fig3-12) dendritic cells efficiently, though increased infection titres appeared to provoke increased DC death (Fig3-11). It was then observed that while HAdV-D30K is highly homologous to HAdV-D49K in terms of amino acid sequence, it displays markedly different infectivity characteristics and does not exhibit a broad tropism as observed in HAdV-C5/D49K (Fig3-13). Exploration of the crystal structures of HAdV-D30K (Fig3-14) and HAdV-D49K (Fig3-15) demonstrated that the residues which are different have their loops exposed to the solvent at the apex of the fiber knob (Fig3-16), and have a large effect on the surface electrostatic profile making HAdV-D49K much more basic (Fig3-17). Finally, a SILAC pull down experiment implicated the desmosome in HAdV-D49K binding to the surface of SKOV-3 cells (Fig-3-18).

3.8.1 Adenovirus diversity is under-represented by whole virus phylogenetics

The investigation of the adenovirus phylogeny showed the expected division of adenoviruses into seven species, A-G (Fig3-1A,B). However, generating the phylogenetic tree using the whole adenovirus genome or by the fiber knob amino acid sequence, gave a different degree of branching within the species. When the tree was generated using the fiber knob sequences the number of nodes within the species D increased compared to the tree generated using the whole adenovirus genomes. This suggests that the degree of diversity within the species D adenovirus fiber knobs may be underestimated based on standard whole genome phylogenies, with more diversity appearing within the fiber knob sequence compared to the rest of the genome. This diversity is likely derived from recombination events between the species D, as has previously been observed, though would require additional analysis to confirm^{32,355}.

Interestingly, the species B adenoviruses were found to display the opposite trend, with less nodes when their phylogeny was analysed based on the fiber knob, suggesting less diversity within the fiber knob proteins compared to other regions of the species B adenovirus genomes. The species B adenoviruses have previously been shown to engage one, or both, of two receptors: CD46 and Desmoglein 2 (DSG2)^{55,89,125}. CD46 and DSG2 are broadly expressed cell surface receptors in humans. It could be that this broad tropism has encouraged development of selectivity post-cell entry, resulting in the greater diversity in other regions of the genome while the entry mechanism is conserved.

The development of the oncolytic adenovirus Enadenotucirev (formerly ColoAd1) seems to support this idea. Enadenotucirev was developed by facilitating recombination between various adenoviruses⁹⁷. The final Enadenotucirev virus is predominantly a recombination between HAdV-B3 and HAdV-B11p, with the capsid proteins being entirely derived from HAdV-B11p, which naturally engages both CD46 and DSG2 as primary cell entry receptors. The recombination with HAdV-B3 imparts mutations in non-capsid protein producing, early genes^{55,89}. DSG2 and CD46 utilisation is retained in the Enadenotucirev virus while proteins expressed post cell entry are thought to be the determinants of cancer cell selectivity¹⁰⁰. Regardless of the selective

pressures which have been exerted on the adenoviruses to recombine in a form which replicates preferentially in colorectal cancer cells, enadenotucirev has maintained the CD46 and DSG2 pan-tropism of HAdV-B11p. Despite the selective pressure to evolve preferentially in colorectal cancer cells, and CD46/DSG2 being non-specific for these cells, the virus retained this receptor tropism. This suggests that the receptor usage of the fiber protein and its receptor prevalence is a stronger determinant, and evolutionary rate limiting step in adenovirus evolution, than the early genes, making it difficult to select for new cellular tropisms in the context of whole adenovirus serotype recombination.

This phylogenetic investigation has exposed hidden sources of adenovirus genetic diversity, in this case within the fiber. Extending this method to compare the diversity within other adenovirus proteins will likely reveal other regions of both increased and reduced variation which may help in predicting the behaviour of previously under studied adenoviruses and elucidate as yet unknown adenovirus-host interactions.

3.8.2 Infectivity studies using HAdV-C5 pseudotypes suggest previously unknown receptor interactions within the species D adenoviruses

In order to investigate the effect of the fiber knob proteins, responsible for the primary adenovirus-cell interaction, we generated pseudotyped of HAdV-C5 vector expressing the fiber knob of species D adenoviruses. These pseudotypes enable us to investigate the effects of the fiber knob proteins and control for the non-fiber knob protein effects. HAdV-C5 (CAR binding) pseudotyped with the fiber knob domains of HAdV-B35 (CD46 binding), HAdV-D10, D26, D30, D48, or D49 were tested in CHO-CAR cells (Fig3-2A). The CAR utilising HAdV-C5 infected CHO-CAR cells efficiently, as expected, while the CD46 utilising HAdV-C5/B35K control did not. The HAdV-D26K, D48K, and D49K all infected approximately 70% of cells, while the HAdV-D30K pseudotype was less efficient infecting 40% of CHO-CAR cells. This was a surprising result since none of these viruses had previously been reported to engage CAR as a primary receptor. This has now been explored in more detail as will be detailed in subsequent chapters (Chapter 4), and is described for HAdV-D26K and

HAdV-D48K in our recent publication⁷². In summary, it appears likely that many of the species D adenovirus fiber knobs possess low affinity for CAR, and that this affinity may be moderated by the length of the DG-loop of the fiber protein and, in some cases, masked by other receptor interactions such as sialic acid³⁵⁶.

We next tested the same pseudotypes in CHO-BC1 cells. The CAR engaging HAdV-C5 was no longer able to infect, while the B35K pseudotype was highly efficient, commensurate with the known HAdV-B35 tropism for CD46 (Fig3-2B)⁸⁹. HAdV-C5/D26K infected approximately 30% of CHO-BC1 cells, which is above the background observed with HAdV-C5. The mechanism by which HAdV-D26K mediates this above background infection is unclear since it has now been shown that HAdV-D26K does not engage CD46 as a primary entry receptor as detailed in the next chapter and our recent publication⁷². It is possible that this is derived from the fact that CD46 is a glycoprotein and HAdV-C5/D26K is utilising sialylated glycosylation motifs on the CD46 to enable infection, as supported by our other recent publication and the work in chapter 5³⁵⁶. However, the level of infectivity observed using HAdV-C5/D49K was significantly higher, matching the 95% of cells infected by the HAdV-C5/B35K pseudotype (Fig3-2B).

It would be tempting, therefore, to assume that HAdV-D49K binds to CD46 as a means of infection, as has previously been proposed³⁵¹. However, similar levels of infectivity (90%) were observed in CHO-K1 cells which do not express CD46 or any other known adenovirus receptors (Fig3-2C), and surface plasmon resonance experiments confirm a lack of any interaction between HAdV-D49K and CD46 (See chapter 4).

We observed that recombinant HAdV-D49K protein was able to inhibit HAdV-C5 infection of CHO-CAR cells (Fig3-5B), and therefore must be able to engage CAR in order to mediate this inhibition. However, it was surprising to see that the same protein was unable to inhibit the HAdV-C5/D49K virus, which uses the same fiber protein, in the CHO-CAR cells (Fig3-5D), though it did inhibit infectivity by nearly 50% in CHO-K1 cells (Fig3-6D). Therefore, HAdV-D49K must be capable of binding to CAR, but with a weaker affinity than HAdV-C5K, since it does not inhibit HAdV-C5 as strongly as the HAdV-C5K (Fig3-5A,B).

This inference was confirmed by surface plasmon resonance and competition inhibition assays, as described in chapter 4. However, it must also be true that transduction by the HAdV-C5/D49K virus is not dependent upon CAR engagement (Fig3-5C, Fig3-6C,D). This is not the only example of a species D adenovirus having a dual tropism. HAdV-D26 is known to utilise CAR, but at low affinity, and to also use sialic acid as a primary cell entry receptor and HAdV-B11p is known to engage both DSG2 and CD46 as a primary cell entry receptor^{55,72,101,356}.

Taken together it seems likely that HAdV-D49K binds to both CAR and another as yet unidentified receptor. Given that HAdV-D49K only appears to be capable of self-inhibition in the absence of CAR (Fig3-6D) and its infection efficiency is reduced in the presence of CAR, we suggest that the differences are due to either post cell-entry factors, or alternative trafficking. One hypothesis is that HAdV-D49K traffics through the cell differently when utilising CAR as a receptor or the other unknown molecule, which should be investigated by intracellular tracking studies.

3.8.3 Structural comparison reveals electrostatic surface profile is key to the broad tropism of HAdV-D49K

Further information about the nature of the broad HAdV-D49K tropism became apparent with the realisation that it differed with HAdV-D30K at only 4 amino acid positions (Fig3-13D). Despite this high degree of homology between HAdV-D49K and HAdV-D30K, the HAdV-D30K pseudotyped virus demonstrated a remarkably different infectivity profile compared to HAdV-C5/D49K, exhibiting only basal level of infection (Fig3-13A,C). We therefore sought to evaluate a structural basis for this biological observation by generating crystal structures of these fiber knob proteins. We observed that one of the amino acid changes at position 337 is likely to be functionally homologous and is not exposed to the solvent (Fig3-13D, Fig3-16A,B). Since this residue is buried in the protein and not capable of forming intermolecular interactions, we presumed that it is not involved in the difference between the HAdV-D30K and D49K receptor binding capabilities.

In contrast, the other three amino acids which differ between HAdV-D30K and D49K are positioned at the turn of two loops on the apex of the fiber knobs

(Fig3-16). The side chains of these residues are solvent exposed; thus they affect the surface properties of the protein. It follows then that these surface changes must be the factor which alters the infectivity characteristics of the virus.

We observed that these three amino acid changes have a significant affect upon the pI of these proteins, making the electrostatic surface of HAdV-D49K more basic than that of HAdV-D30K (Fig3-17). The substitution of glycine in HAdV-D30K for alanine in HAdV-D49K at position 330 (Fig3-13D) adds a small hydrophobic motif, the effect of which is difficult to speculate on. It is the E-K and Q-K substitutions at positions 238 and 331 in D30K and D49K, respectively, drive the change in surface potential. Superposition of the two crystal structures and comparison of the altered residues shows that while the substitutions alter the electrostatic profile they are similar in terms of their orientation and steric bulk (Fig3-16C,D). It should be noted that these surfaces are a snapshot of the average surface conditions seen in the crystal lattice. Solvent exposed side chains of this length are likely to be flexible under biological conditions, creating a mobile and constantly fluctuating electrostatic surface.

The central apical depression of HAdV-D30K is seen to be moderately basic (Fig3-17). This depression lays between three apical protrusions formed of the core fiber knob β -strands and the loops pictured in figures 3-16C and 3-16D. As the viewer looks around the central depression climbing the protrusions toward the loop the electrostatic surface becomes more neutral and eventually slightly positive at the tip of the protrusions: E238. The three amino acid substitution disrupt this gradient in HAdV-D49K where the protrusions remain electropositive, culminating in the apical K238. This means that HAdV-D49K has an electropositive funnel leading to the central depression, though the significance of this is unclear it would seem plausible that this funnel may guide an oppositely charged molecule towards a binding site in the central depression.

3.8.4 Pull down implicates the desmosome in HAdV-D49K mediated infection

In an attempt to isolate the cell surface receptor, we performed the SILAC Co-precipitation experiment in figure 3-18. It was therefore disappointing that the

experiment did not isolate any cell surface proteins, however it may still provide valuable insights into the potential receptor, or the pathway(s) upon which viral entry depends.

The pull down of annexin and GAPDH could potentially be explained by their roles in vesicle trafficking. Annexin has previously been shown to be involved in the trafficking of membrane binding proteins, including viruses, to the golgi and ER³⁵⁷⁻³⁵⁹. GAPDH has been shown to be involved in vesicle transport between the golgi apparatus and the ER³⁶⁰. The reason the pull down of CLK3 is less clear, but could be due to its regulation of mRNA splicing³⁶¹. It would make sense for all these proteins to be upregulated in response to a large influx of vesicles endocytosed from the cell surface, which is a likely result of binding large numbers of foreign proteins to receptors on the cell surface such as HAdV-D49K.

The discovery that DSP and JUP were pulled down in association with HAdV-D49K was particularly interesting (Fig3-18). Both JUP and DSP are involved in the formation of desmosomes. DSP has been shown to interact with JUP and bridge its interaction to components of the cytoskeleton, while JUP can form interactions with numerous cadherin proteins, including DSG2, a known adenovirus receptor^{354,362,363}.

While desmoglein 2 is known to interact with species B adenoviruses, it has not previously been observed to interact with species D adenoviruses^{54,55,125}. In the next chapter we further investigate the potential for interaction between HAdV-D49K and DSG2 by surface plasmon resonance and did not detect any binding, while we confirm the previously reported interaction between HAdV-B3K and DSG2 (See chapter 4). However, if the desmosome is involved with HAdV-D49K surface binding there are numerous other partners for JUP which are structurally homologous to DSG2, other members of the cadherin family, and are therefore potential HAdV-D49K partners which should be examined in future studies.

If HAdV-D49K does bind to a cadherin molecule, in a manner similar to how HAdV-B3K interacts with DSG2, this could explain why they were not isolated by the SILAC Co-precipitation experiment¹²⁵. HAdV-D3K has been shown to preferentially form a 1:1 stoichiometry with DSG2 due to steric hinderance, and

to require multimerization to efficiently infect cells, which has been suggested to be a driving force behind the viruses generation of penton dodecahedrons^{54,108,110,125}. Our own studies in chapter 4 confirm that HAdV-B3K binding to DSG2 is a relatively low affinity interaction. SILAC Co-precipitation assumes a stable interaction between ligand and receptor. However, if this complex was unstable then the complex could break down and reform during wash and incubation steps during the lysis and Co-precipitation steps of the experiment, which would equilibrate the heavy to light peptide ratio as reformed complexes could come from either cell population after sample mixing. Further, the SILAC Co-precipitation experiment was performed using recombinant HAdV-D49K fiber knob which would not reveal interactions which require polyvalence, presumably as the low affinity necessitates this for a stronger avidity, such as the HAdV-B3K interaction with DSG2.

Further SILAC Co-precipitation experiments should be modified to account for these limitations by utilising crosslinking prior to the mixing of the heavy and light labelled samples to prevent loss of low affinity interactions. Furthermore, the SILAC Co-precipitation experiment could be repeated using whole virus to grant polyvalence, enhanced avidity, and hopefully improve the quantity of receptor in the initial pulldown.

The final limitation of SILAC Co-precipitation is that it assumes the proteins primary binding partner is another protein. However, it is known that many adenoviruses, including members of the species D, utilise sialylated glycans as binding partners, and may interact with heparan sulphate proteoglycans (HSPGs) either as receptors or as a cellular defence mechanisms (decoy receptors)^{27,356,364,365}. If this is the case the SILAC Co-precipitation will not pull-down specific proteins, but instead would be likely to co-precipitate with peptides associated with glycosylation motifs on the myriad glycoproteins on the cell surface, assuming they are not lost due to the aforementioned issues with low affinity. It would be interesting to retool the library screening software to reanalyse the mass spectrometry outputs and quantify the ratio of peptides in which consensus glycosylation motifs are located. Were the HAdV-D49K interacting with a glycan, as seems plausible given the results in chapter 5, we would expect to see a statistically significant increase in pulled down peptides

containing consensus glycosylation motifs in the heavy isotope labelled cells incubated with HAdV-D49K.

3.8.5 Infectivity studies of HAdV-C5/D49K in cancer cell lines reveal high efficiency infection and pan-tropism

Given the broad infectivity profile observed using HAdV-D49K we sought to explore its utility as a vector in cancer cell lines, with a view to using it as a virotherapy. Testing in nine pancreatic cancer cell lines showed HAdV-C5/D49K infecting the cells more efficiently than HAdV-C5 in all instances (Fig3-7). However, the degree to which the efficiency of infection improved was highly variable, ranging from a 4.6x improvement in MIA PaCa2 cells to a 211x improvement in BxPc3 cells, even though both lines are pancreatic adenocarcinoma. A similar pattern was observed in five breast cancer cell lines where, again, all tested lines were more efficiently transduced by HAdV-C5/D49K than HAdV-C5 (Fig3-8). In this instance the change in efficiency ranged from a 4.6x improvement in MDA-MB-231 cells to a 497x improvement in MCF7 cells, both cell lines are breast adenocarcinoma cell lines derived from metastasis found in pleural effusions. Finally, the HAdV-C5/D49K pseudotype was tested in three lung cancer cell lines, an oesophageal, colorectal, and ovarian cancer cell line (Fig3-9). The A427 lung cancer cell line and DLD-1 colorectal cancer cell lines were the only ones of the 20 tested lines where HAdV-C5/D49K was not more efficient at infecting the cells.

Looking at the relative level of transduction observed in the lines, as measured by the luciferase transduction assay, HAdV-C5/D49K appears to transduce all lines effectively. Rather than the variation in fold change between the two vectors being due to lack of transformation by HAdV-C5/D49K, it is due to improvement in transduction by HAdV-C5. For example, in MCF-7 cells HAdV-C5 infection produces $<1 \times 10^6$ RLU/mg while HAdV-C5/D49K produces $>6 \times 10^6$ RLU/mg (Fig3-8). In DLD-1 cells the fold change is not significant between the two vectors, but HAdV-C5/D49K still mediates 5×10^6 RLU/mg, it is just that HAdV-C5 is also producing 6×10^6 RLU/mg. It therefore seems that the HAdV-D49 fiber knob grants an apparent pan tropism, with none of the tested cell lines refractory to infection.

3.8.6 HAdV-C5/D49K can infect dendritic cells with high efficiency and may have utility as a vaccine vector

Given this pan tropism, and without a known mechanism by which the HAdV-D49K grants this ability, it is difficult to engineer a more refined tropism by which this virus could be used as a tumour targeted virotherapy with minimal off target infection. We therefore considered alternative applications whereby pan tropism, and especially highly efficient cell transformation, could be beneficial. Other work with adenovirus vectors has highlighted species D adenoviruses as potential viral vaccine vectors, such as the HAdV-D26.ZEBOV vaccine, antigens for more pathogenic viruses are encoded within and expressed by the adenovirus after infecting the patient. Immune cells recruited to the site of infection are then presented with the pathogen antigen with the goal of generating a persistent immune response^{36,40,353}. Dendritic cells are particularly important in this response as broad range antigen presenting cell capable of stimulating T-cell response and memory formation. Therefore antigen “loading” of DC can mediate stronger immune responses capable of generating lasting humoral immunity and to break tumour immune tolerance^{353,366,367}.

We therefore tested the ability of HAdV-D49K to infect dendritic cells. We tested the ability of HAdV-C5/D49K to infect human dendritic cells derived from donor blood monocytes (human monocyte derived dendritic cells – hMDDC). We observed that HAdV-C5/D49K was significantly more efficient than HAdV-C5, achieving nearly 90% infection in the first experiment (Fig3-10). A repeat of this experiment using cells from a different donor also showed HAdV-C5/D49K was more efficient than HAdV-C5, though the increase seen when infecting with AHdV-C5/D49K was less (Fig3-11). Comparing virus given at the same dose of 1000 VP/Cell it appears that this is not due to less efficient infection by HAdV-C5/D49K in the second donor, but due to more efficient infection by HAdV-C5. In this experiment we tried a broader dosage of virus and stained for only the live population. We observed that while higher doses resulted in a greater proportion of the cells being infected there were fewer cells in the live gate as the virus dose increases, and less cells in the live gate for HAdV-C5/D49K than for HAdV-C5 at equivalent doses (Fig3-11). It seems clear that increased viral transduction of DC results in greater cell death, therefore the HAdV-C5/D49K

virus, while efficient at infecting the hMDDC, also results in more death. However, the cause of this death is not clear, since the viruses used in this infectivity experiment are replication incompetent and thus death is not a result of adenovirus mediated lysis. We can speculate that the mechanism of hMDDC death may be due to overstimulation of the cells. This has previously been reported as a cause of hMDDC death and proposed as a means of regulating the magnitude and of, and maintaining, immune tolerance³⁶⁸. However, whilst this is an important proof of concept experiment demonstrating the potential of this vector to infect hMDDC and express foreign proteins, isolated hMDDC expressing GFP are of little relevance therapeutically.

Therefore, to further investigate the potential of this virus as a vaccine vector requires *in vivo* studies. We tested the HAdV-C5/D49K virus in mouse MDDC (mMDDC) to establish whether similar infectivity patterns are seen and therefore the potential to run further studies of the vector in mice, an experiment which is ongoing. We observed similar results to the hMDDC study in so far as the mMDDC was infected approximately 40% more efficiently than the HAdV-C5 treated mMDDC and less live cells at increased viral doses (Fig3-12). The only difference we observe is that, for a given dose, the mMDDC are infected more efficiently than hMDDC. Based on these results we have generated versions of the HAdV-C5/D49K virus expressing cancer (5T4 and NY-ESO) and viral (pp65) antigens and will be continue to assess the value of this virus as a vaccine vector in mouse models of disease.

Based on the results in this chapter we cannot definitively identify the receptor binding partner of HAdV-D49K. However, the results are strongly suggestive of a low affinity interaction with an as yet uncharacterised adenovirus receptor. These results and those in the subsequent chapters show that HAdV-D49K can bind to CAR, but that it is not dependent upon it for cell infection (Fig3-2, 3-5, 3-6), a trait which appear to be common in the species D adenoviruses^{72,356}. Our Co-precipitation experiment indicate that the receptor may be associated with the desmosome (Fig3-18) but is unlikely to be DSG2 (see chapter 4).

Considering these results regarding the receptor tropism of HAdV-D49K it is less surprising that we see such a broad cell tropism, and ability to infect so many different cell lines efficiently (Fig3-7-12). The ability to efficiently

transduce DC is especially of interest given the recent success of the HAdV-D26 vector as a vaccine^{40,105,341,369}. Further studies of this vector should account for this broad tropism, and focus upon applications in which this is beneficial, being mindful of the likelihood of off target interactions.

3.9 Summary of Chapter 3

The major findings in this chapter have been:

1. Different adenovirus species appear to have differing levels of diversity localised to particular regions of their genomes. This is evidenced in Figure 3-1, especially figure 3-1C which shows the regions of variability.
 - a. Species D adenoviruses have more diverse fiber knob domains than species B adenoviruses. Figure 3-1A and B demonstrates this by showing the number of separate nodes generated during phylogenetic analysis by fiber knob amino acid sequence or whole viral genome, respectively.
2. HAdV-C5/D49K can infect a broad range of cells with very high transduction efficiency.
 - a. Cancer cell lines (Fig3-1 to 3-9).
 - b. Dendritic cells (Fig3-10 to 3-12).
3. HAdV-D49K differs from HAdV-D30K in only four amino acids, 3 of which are exposed, but HAdV-D30K does not exhibit the same broad tropism as HAdV-D49K, as seen in cell infectivity assays (Fig3-13).
 - a. These changes result in radical alterations to the electrostatic surface potential. This is seen in the electrostatic surface potential of the two proteins, as compared in Figure 3-17.
4. SILAC Co-precipitation experiments indicate that the desmosome may be important in HAdV-D49K mediated cell infection. This is evidenced by the pull down of two desmosome associated proteins, desmoplakin and desmoglobin (Fig3-18).

CHAPTER 4. Results: Characterisation of three species D adenovirus fiber knob proteins and their interaction with known adenovirus receptors

4.1 Introduction:

Designing efficacious viral vectors requires intricate knowledge of the underlying biology of the virus from which the vector is derived¹⁴. This essential knowledge is a pre-requisite to manipulate the vector into behaving as a therapeutic agent, rather than as a pathogen.

Inappropriate consideration and manipulation of the virus from which the vector is derived will result in only a small proportion of vector successfully reaching the intended tissue, with the majority being sequestered in off-target tissues or neutralised by antiviral immune responses. This will result in a larger dose being required to deliver enough vector to the diseased tissue to have a therapeutic effect. Such high doses of virus increase the risk of significant dose limiting toxicities and associated adverse immune responses, such as cytokine release syndrome. Whilst vectors can be engineered to replicate selectively post infection, it is clearly preferable, where possible, to achieve equivalent therapeutic index using a smaller dose of virus to alleviate dose limiting toxicities.

Altering viral vector tropism towards a specific site of disease requires defining the native means of cellular entry and infection. This requires detailed knowledge of the native viral receptor, the mechanism by which the virus binds these receptors, in order to make the refinements necessary to ablate such interactions. This reduces binding and sequestration of the virus in non-target tissues, thus increasing the pool of vector available to be trafficked to sites of disease²⁶².

The adenovirus fiber knob protein is the primary tropism determinant of adenovirus, as it is the classical means by which the virus attaches to the cells and begins the infection pathway^{64,123}. Non-classical tropism determinants are indirect cell surface attachment, whereby a protein interacts with the adenovirus capsid and another cell surface protein to initiate infection. An example of this non-classical pathway is the binding of coagulation factor X (FX) to the hexon and then to heparan sulphate proteoglycans (HSPG)^{63,168,169}. In this chapter we provide structural insight into the fiber knob proteins of 3 clinically important species D adenoviruses; Human adenovirus species D type 26 fiber knob (HAdV-D26K), Human adenovirus species D type 48 fiber knob (HAdV-D48K),

and Human adenovirus species D type 49 fiber knob (HAdV-D49K), and investigate their ability to bind to previously identified adenovirus receptors CD46 (MCP, Membrane Cofactor Protein)^{51,92,370,371}, Desmoglein 2 (DSG2)^{54,55,110,125}, and coxsackie and adenovirus receptor (CAR)^{14,75,372}.

To achieve this, we developed an integrative structural biology workflow based on newly determined crystal structures, and confirmed the predictions using competition inhibition assays and surface plasmon resonance. Using this strategy, we clarify the receptor usage of these species D adenoviruses by investigating their affinity for known receptors.

4.2 Novel species D fiber knob crystal structures:

4.2.1 Expression and purification of recombinant fiber knob proteins

After expression in SG13009 *E.coli* we lysed the harvested bacteria and purified the lysate by Ni-IMAC (Fig4-1A). We observed a minor peak of unknown content, followed by a large major peak beginning at 50% concentration of the imidazole containing buffer, equivalent to 200mM imidazole (Fig4-1A). We assessed the purity of these fractions by SDS-PAGE and Coomassie staining, demonstrating that with careful fraction collection Ni-IMAC purification alone may be enough to achieve high purity recombinant fiber knob protein (Fig4-1B). However, in some fractions small fragments are visible below 25kDa, indicating possible impurities. The major protein bands ran around 27kDa when boiled and reduced, the size expected of a monomeric fiber knob domain. The unreduced and boiled fractions show a major band at around 68kDa, below the expected size of the trimer, though above the expected size of dimer.

The second purification step, size exclusion chromatography (SEC), produced a very sharp major peak prior to a similar peak in the conductance, demonstrating that the protein has exchanged into the SEC running buffer (Fig4-1C). Analysis of the fractions (Fig4-1D) shows extremely high protein purity, and that the protein runs at the expected size of around 75kDa when run natively, indicative of a trimer. Interestingly some small, approximately 15kDa, protein bands are still visible in the higher concentration fractions, as were seen in the previous purification step. These bands persisted even after repeating SEC. Since all proteins in a given fraction should be of a similar size these smaller fragments may be generated during the running of the protein on the gel rather than during purification. Matching smaller “ghost” bands are seen below the trimeric protein for the same fractions. It is possible the smallest fragments are cleaved tail regions from the fiber knob, consisting of the His-tag and shaft fragment, while the ghost trimer bands and full trimeric fiber knobs in which the tails have been lost, affecting the running on the gel.

These proteins were used for several assays, including crystallisation, calculation of IC_{50} in competition inhibition assays, and surface plasmon resonance (SPR).

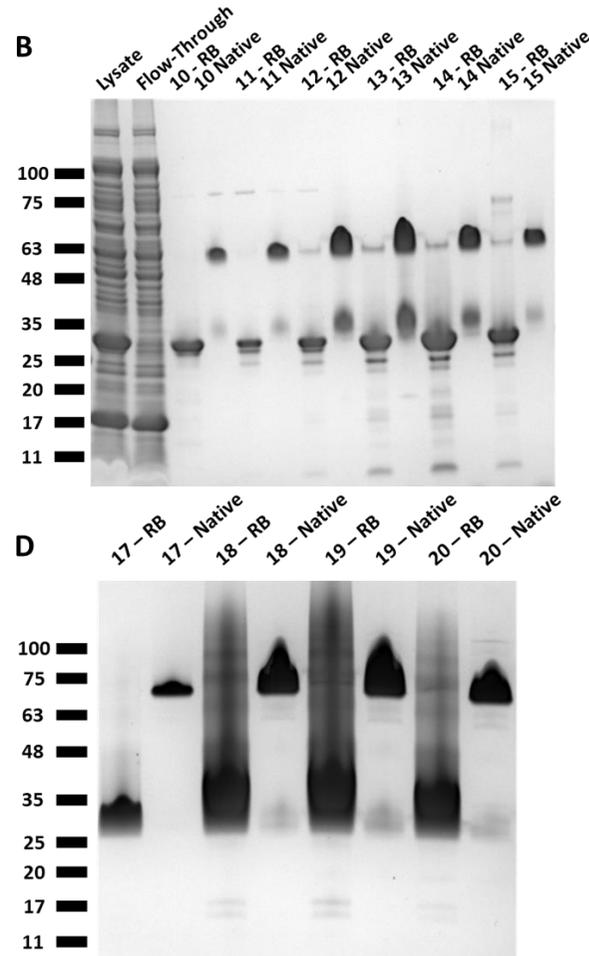
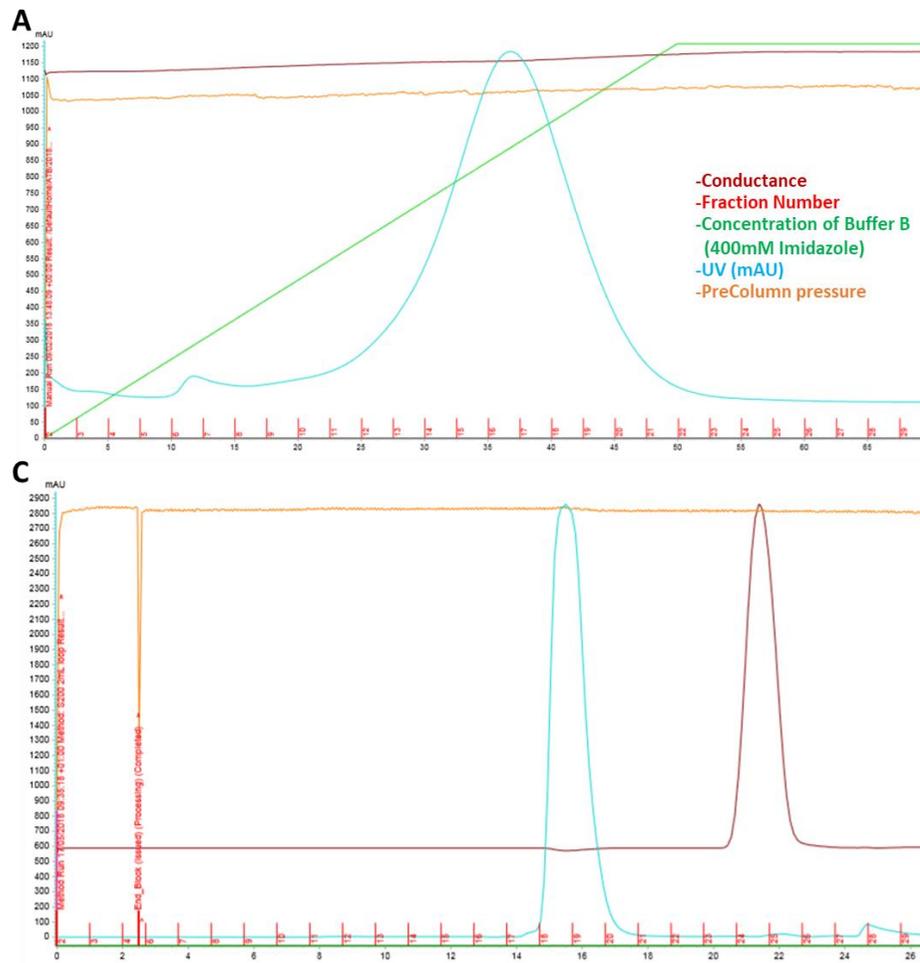


Figure 4-1 Example of purification of fiber-knob protein by Ni-IMAC and SEC. Chromatography trace of fiber-knob protein elution from the Ni-IMAC column by gradient elution shows a minor peak early in the elution (fractions 6 and 7) and a major peak spanning fractions 11-23 (A). SDS-PAGE gels stained with Coomassie brilliant blue reveal the analysed fractions, run both native and reduced with 0.1M DTT and boiled (RB), run around 65kDa and 27kDa, respectively, with other small fractions also visible (B). Purification of the fractions from the major peak of the Ni-IMAC elution by SEC gives a sharp elution peak between fractions 17 and 20, with a small peak visible at fraction 28 (C). SDS-PAGE and coomassie brilliant blue staining of the major peak fractions reveal protein at around 29kDa when run reduced and boiled, and at around 72kDa when run native, corresponding with the predicted monomeric and trimeric molecular weights expected from an adenovirus fiber-knob protein (D).

4.2.2 Novel crystal structures of human adenovirus fiber knob proteins:

The novel crystal structures of 3 fiber knob proteins crystallised in this study, as well as a new, higher resolution structure for adenovirus serotype 5 fiber knob protein are seen in figures 4-2 to 4-5. Different fiber knobs appeared in different space groups with variable resolutions. All are deposited in the wwPDB, validated, and authorised for release. Statistics for the structures solved and deposited in this study are given in Table 4-1.

The asymmetric unit of HAdV-D26K contains a single fiber knob monomer (Fig4-2A). The exceptionally high resolution obtained (0.97Å) meant clear density was visible for numerous water molecules. No other non-protein molecules were evident in the density. The map showed the high level of detail expected of a structure at this resolution (Fig4-2B). The summary PDB report (Fig4-2C) shows the key measures of quality, as determined in the deposition process. While the R_{free} value is excellent compared to most structures it is not as high as many structures at similar resolution, though the data is an excellent fit to the density. This may be due to the lack of anisotropic refinement during structure refinement prior to deposition.

HAdV-D48K crystallised in a different space group to HAdV-D26K, resulting in a larger asymmetric unit than HAdV-D26K, containing two opposing monomers (Fig4-3A). Translation of the asymmetric unit creates a pair of trimers with unbroken 3-fold symmetry, generating two separate examples of the biological unit. HAdV-D48K is characterised by several very long loops, found coordinated to sulphate ions, as well as glycerol and 1,2-Ethanediol (EDO). This coordination does not completely alleviate the flexibility inherent to these long loops, resulting in a lower resolution map (Fig4-3B), and correspondingly weaker statistics (Fig4-3C, Table.1), but the map still shows clear density for side chains enabling the determination of a complete structure.

The existing structure of HAdV-C5K was generated by Xia et al in 1995 (PDB 1KNB)³⁷³ and has no associated map data. We generated a new structure for HAdV-C5K at higher resolution (1.4Å, compared to 1.7Å), which solved as a single trimer copy per asymmetric unit (Fig4-4A), with an Mg²⁺ ion at the apex

of the fiber knob on the 3-fold axis and several EDO molecules contacting loops.

Table 4-1 Statistics of the adenovirus fiber-knob crystal structures generated in this study, as deposited in the wwPDB. 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. MMT (DL-Malic acid, MES monohydrate, Tris: pH 4.0-9.0).

PDB Entry	6FJN – Ad26K	6HCN – Ad5K	6FJQ – Ad48K
Diamond Beamline	I04	I24	I04
Date	2017-05-12	2018-01-26	2017-05-12
Wavelength	0.9795	0.96859	0.9795
Crystallisation Conditions	0.1 M MMT 25% PEG 1500	0.1 M MMT, 25% PEG 1500	0.1 M Bis-Tris- propane, 20% PEG 3350, 0.2M NaNO ₃
pH	6.0	7.0	6.5
<i>a,b,c</i> (Å)	86.01,86.01,86.01	102.16,102.44,77.0	145.18,145.18,145.18
α,β,γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Space group	P 2 ₁ 3	P 2 ₁ 2 ₁ 2	P 4 ₃ 3 2
Resolution (Å)	0.97-60.82	1.49-61.56	2.91-83.82
Outer shell	0.97-1.00	1.49-1.53	2.91-2.99
R-merge (%)	4.3 (74.5)	13.4 (183.8)	12.5 (302.6)
R-meas (%)	4.5 (97.5)	15.9 (218.3)	12.7 (306.3)
CC1/2	1.00 (0.427)	0.983 (0.565)	1.00 (0.705)
I / σ (I)	27.3 (0.7)	7.1 (0.7)	22.2 (1.7)
Completeness (%)	94.9 (43.9)	99.8 (99.9)	100.0 (100.0)
Multiplicity	16.7 (1.6)	6.6 (6.3)	41.2 (41.4)
Total Measurements	1,978,768 (6,429)	876,648 (60,950)	496,751 (5,136)
Unique Reflections	118,603 (4,055)	131,951 (9,638)	12,061 (877)
Wilson B-factor(Å ²)	8.2	18.3	74.5
R-work reflections	112,612	125,479	11,371
R-free reflections	5,879	6,388	572
R-work/R-free (%)	18.2 / 19.5	21.1/23.3	20.1 / 29.1
Bond lengths (Å)	0.025	0.011	0.019
Bond Angles (°)	2.339	1.534	2.293
¹ Coordinate error	0.020	0.087	0.370
Mean B value (Å ²)	17.6	30.6	84.9
Favoured/allowed/Outliers	138 / 7 / 1	133 / 10 / 1	350 / 28 / 7
%	94.5 / 4.8 / 0.7	92.4 / 6.9 / 0.7	90.7 / 7.5 / 1.8

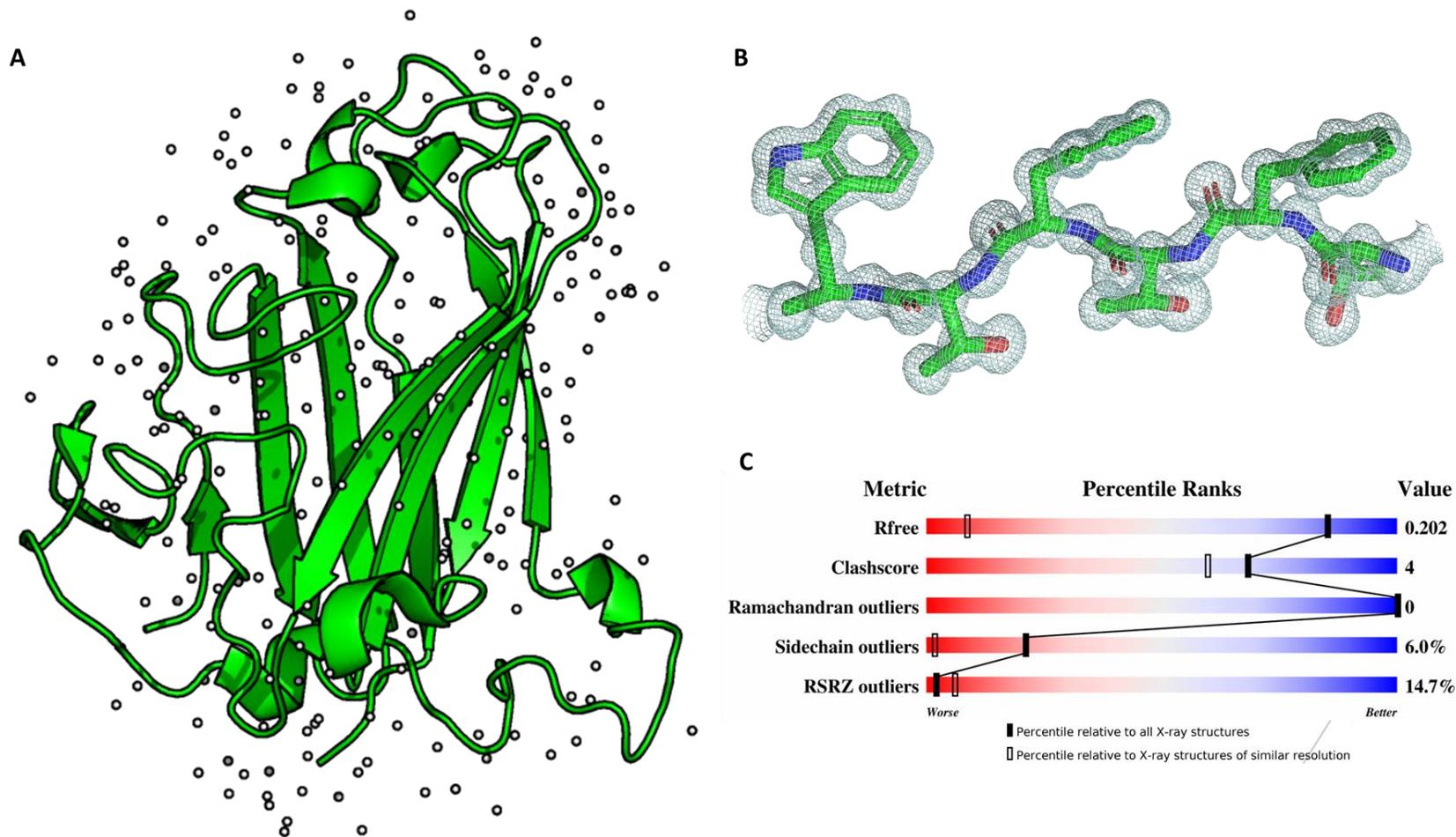


Figure 4-2 Overview of the solved crystal structure of HAdV-D26K, deposited in the wwPDB as 6FJN. The asymmetric unit consisted of a single monomer of the fiber-knob (**A**), the quality of the map (blue mesh) through the main β -strands was excellent and clearly defined the location of each residue (**B**). The statistics showed acceptable scores compared to other wwPDB entries (**C**). Map is displayed at $\sigma=1.5$, and carved at 1.5\AA . Protein is shown as a cartoon representation. Waters are shown as white, non-bonded spheres.

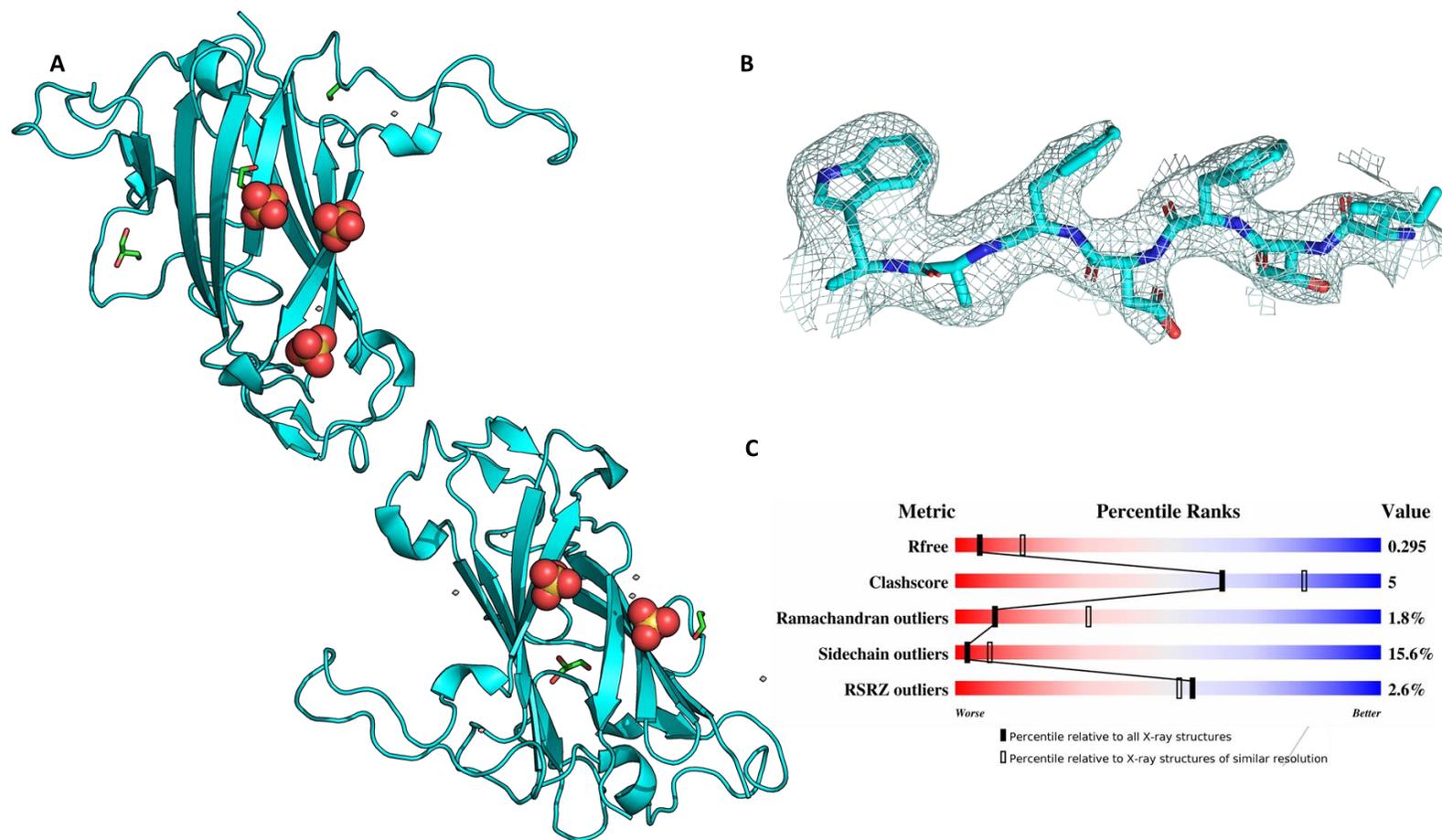


Figure 4-3 Overview of the solved crystal structure of HAdV-D48K, deposited in the wwPDB as 6FJQ. The asymmetric unit consisted of two apically opposed fiber-knob monomers (A), the quality of the map (blue mesh) through the main β -strands enabled definition of individual side chain orientations (B). The statistics showed acceptable scores compared to other wwPDB entries (C). Map is displayed at $\sigma=1.0$, and carved at 2.0\AA . Protein is shown as a cartoon representation. Waters are shown as white, non-bonded spheres. Sulfate ions are shown as a space filling model, with oxygen atoms in red and sulfur in yellow. EDO and Glycerol molecules are seen as a stick representation in green, with oxygen atoms in red.

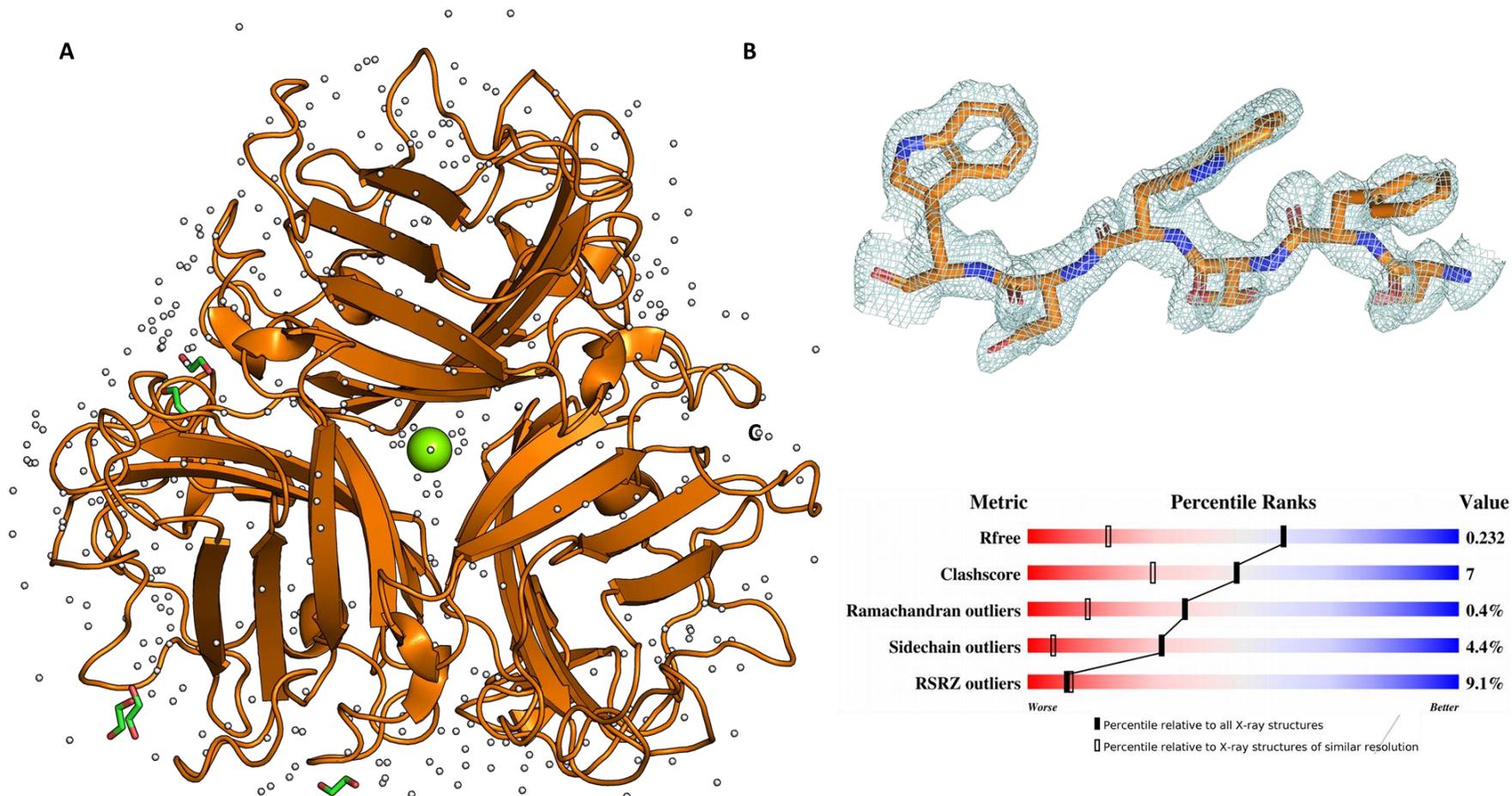


Figure 4-4 Overview of the solved crystal structure of HAdV-C5K, deposited in the wwPDB as 6HCN. The asymmetric unit consisted of 3 fiber-knob monomers, creating a trimer which is also the biological unit (A), the quality of the map (blue mesh) through the main β -strands was excellent and clearly defined the location of each residue (B). The statistics showed acceptable scores compared to other wwPDB entries (C). Map is displayed at $\sigma=1.5$, and carved at 1.5\AA . Protein is shown as a cartoon representation. Waters are shown as white, non-bonded spheres. EDO molecules are seen as a stick representation in green, with oxygen atoms in red

The map clearly defines the location of each residue (Fig4-4B) and obtained good statistics (Fig4-4C). It is notable that while the structures are almost identical (RMSD 0.19Å) the 1KNB structure differs from this one significantly in the orientation of the DG-loop, suggesting flexibility.

By this means we were able to generate reliable models of previously undescribed fiber knob proteins and place the main chain carbons and side chains. This structural data provides a basis for exploration of the interface forming regions through computational analysis and guided the remainder of

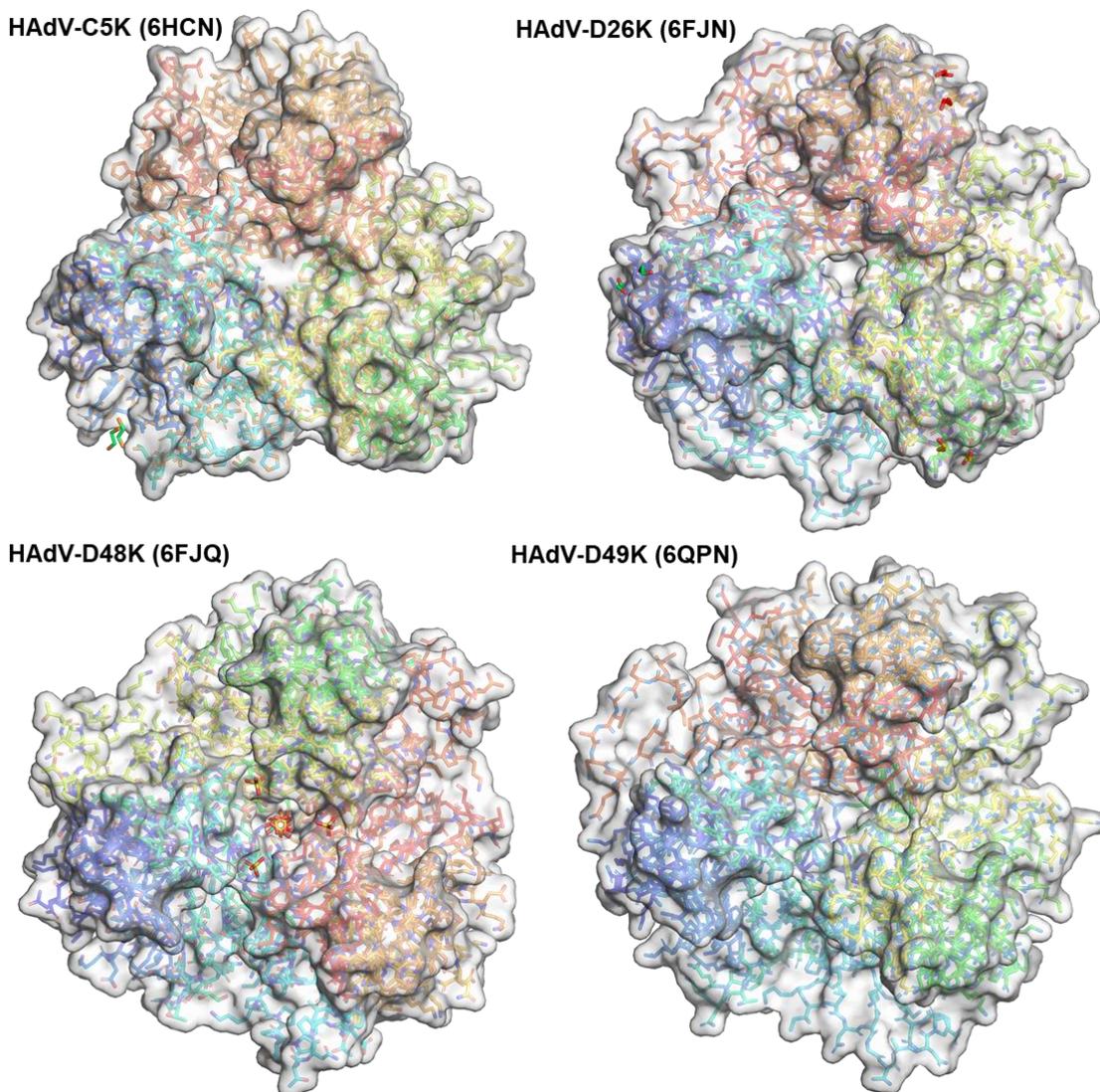


Figure 4-5 Overview of novel crystal structures of species C and D adenoviruses fiber-knob trimers generated in this study. When solved all crystal structures either formed a biological trimer in the native asymmetric unit or enable the reconstruction of the biological trimer through symmetry translations. Representative biological trimers of each adenovirus fiber-knob domain are shown. Individual residues are shown as sticks, coloured as a spectrum from green to blue to red, with the surface representation of the trimer shown as a semi-transparent white surface.

this study.

4.2.3 Biological trimers determined from crystallographic structures:

The biological units obtained from each of the discussed crystal structures are summarised in Fig4-5. The trimeric fiber knob domains of the 4 structures are viewed looking toward the apex of the fiber knob, down the 3-fold symmetry axis: the point most distal from the virus' centre. They are similar in that they all exhibit a central depression, surrounded by extended loops of varying length. They all exhibit 3-fold symmetry, though this is imperfect in structures solved as full trimers rather than individual monomers. The structure has a similar overall fold, having the β -sandwich structure which is prototypical of adenovirus fiber knobs (and many other virus attachment proteins) as first described by Xia et al³⁷³. The differences are in the loops which are highly heterologous, both in the cleft between monomers in the lateral regions, and in the apex extending away

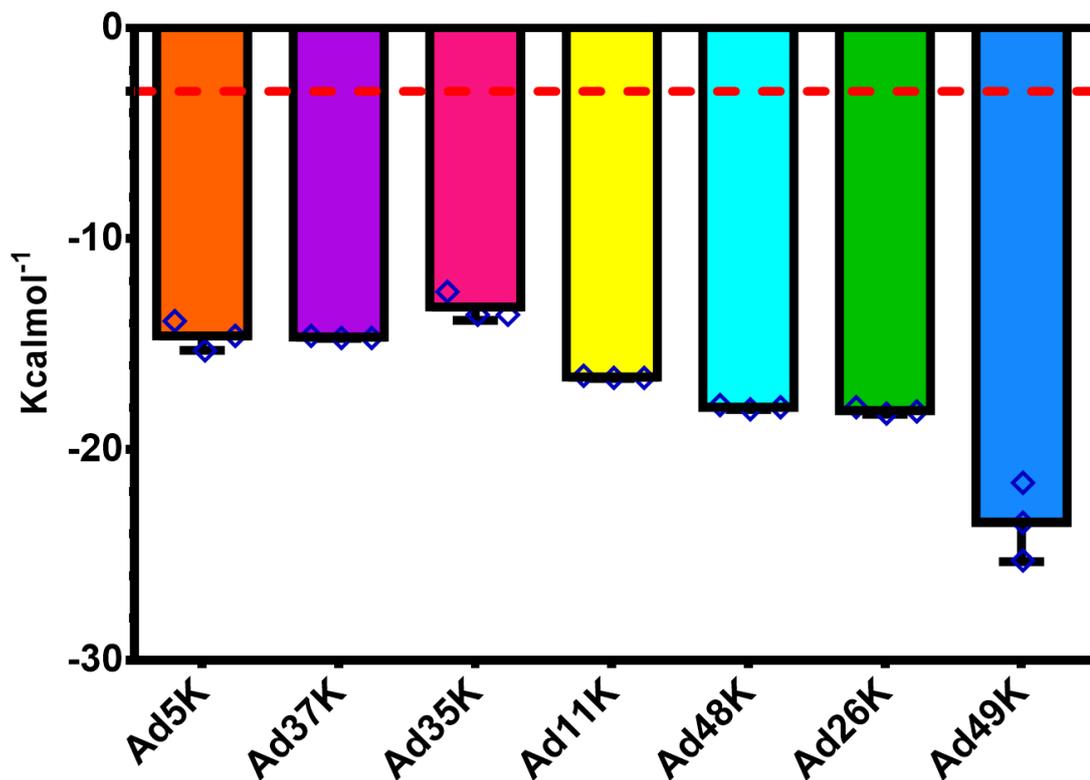


Figure 4-6 Stability of adenovirus fiber-knob trimers. QT-PISA estimates of the interface energy of fiber-knob trimers based upon their crystal structures, suggesting HAdV-D49K may have a substantially stronger trimeric interface than other fiber-knob trimers, whilst HAdV-D48K and HAdV-D26K are the next strongest. Values are the mean of $n=3$ calculations based upon the 3 interfaces found in each native trimer, error is standard deviation from the mean. Dashed red line denotes -3.0 Kcalmol⁻¹.

from the depression.

Despite the conserved main fold, the trimers show variation in the strength of the trimeric interface, as calculated by QT-PISA (Fig4-6). HAdV-D5K and D37K are seen to be very similar in interface binding energy, with HAdV-B35K slightly weaker. HAdV-B11K is slightly stronger than HAdV-D5K and weaker than D48K and D26K. However, HAdV-D49K exhibits a substantially stronger binding energy than any of the other fiber knobs. The significance of this observation is not yet clear, but it follows that it would affect the efficiency of trimer formation.

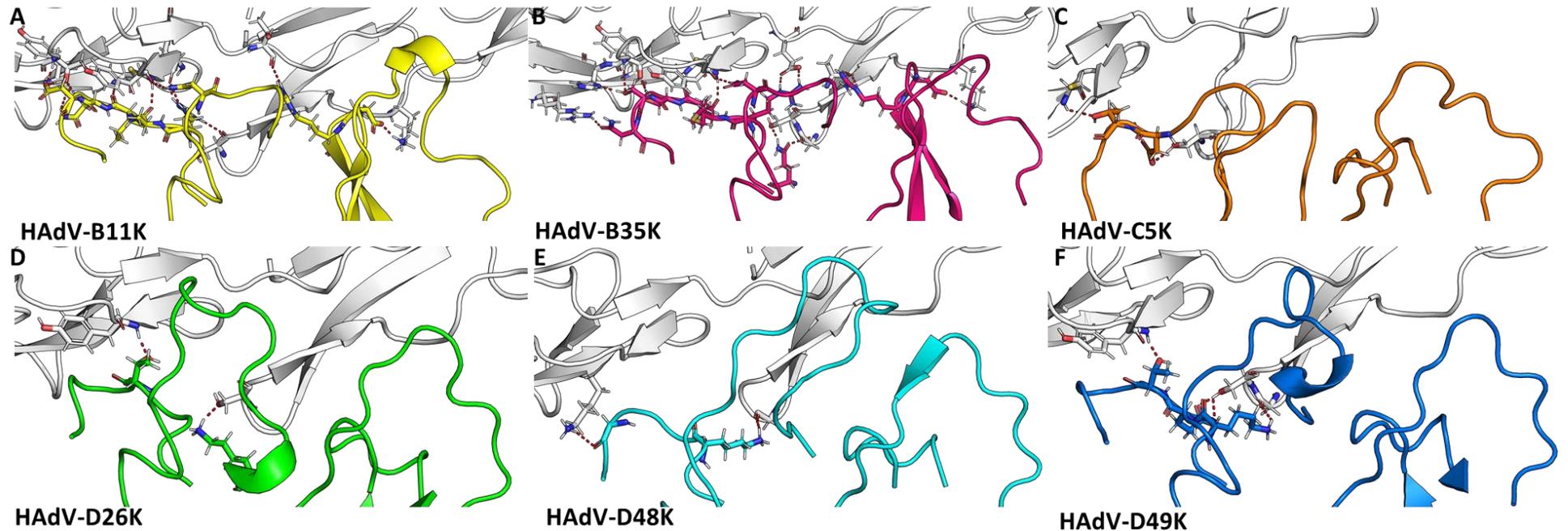
4.3 Examination of species D adenovirus fiber knob binding to known adenovirus receptor proteins:

4.3.1 Investigated species D adenovirus fiber knobs are unlikely to utilise CD46 as a primary receptor:

We utilised the crystal structures of HAdV-D26K, HAdV-D48K, and HAdV-D49K to produce models of their fiber knob proteins in complex with CD46. As well as the template structure, the previously published HAdV-B11K in complex with CD46⁸⁸, we generated similar models using the non-CD46 binding fiber knob of HAdV-C5K, and the CD46 engaging fiber knob of HAdV-B35K as negative and positive controls, respectively. The template PDB of HAdV-B11K in complex was treated with the same modelling procedure.

The CD46 binding interface comprises of CD46 domains SC1 and SC2 and the HI, DG, GH, and IJ loops of HAdV-D11K (Fig4-7A). It was the interface at these loops which we investigated by modelling. Both HAdV-B11K and HAdV-B35K, known CD46 utilising viruses were seen to form numerous polar bond with CD46 (Fig4-7A,B), as would be expected from a strong interface. HAdV-C5K formed very few polar interactions with CD46, again as would be expected from a weak interface by a known non-CD46 binding virus (Fig4-7C). HAdV-D26K formed only 2 polar interactions with CD46 (Fig4-7D), the same as HAdV-C5K, and HAdV-D48K (Fig4-7E). HAdV-D49K formed 3 polar contacts with CD46 (Fig4-7F). The species C and D adenovirus fiber knobs were only capable of forming contacts within the HI loop, possibly owing to broadly similar morphology to the species B in the region.

Amino acid sequence alignment of the modelled fiber knob proteins shows significant sequence divergence between the species B compared to the species C and D fiber knobs (Fig4-7G). The species B, CD46 binding, fiber knobs form more polar contacts than the others are predicted to and that they do so in more



G

HAAdV-C5K	400	TLWTTTPAPS-PNCRINA---EKDAKLTFLVLTCKGSSQILATVSVLAVKGS LAPISCTVQ-----SAHLIIRFDENGVLIN
HAAdV-D26K	188	TLWTTFPDPS-PNCKMST---EKDSKLTFLVLTCKGSSQVLGNVSLLA VTFGEYHQMTATP---KKDVIISLLPDENGVLLP
HAAdV-D48K	182	TLWTTFPDPS-PNCKIDQ---DKDSKLTFLVLTCKGSSQILANMSLLLVVKGKFSMINNKVNGTDDYKFKFTIKLLFDEKGVLLK
HAAdV-D49K	188	TLWTTFPDPS-PNCKVSE---EKDSKLTFLVLTCKGSSQILASVSLLVVKGKFPANINNKVNGTDDYKFKFSVLLFDANGKLLT
HAAdV-B11K	133	TLWTCVNPTEANCQIMNSSESNDCKLILTLVKTGALVTAFAVYVIGVSNFNMLTTHR-----NINFTAELFFDSTGNLLT
HAAdV-B35K	133	TLWTCGINPP-PNCQIVENTNTDNGKLTFLVLVKNGGQLVNGVYSLVGVSDTVNQMFPTQK-----TANIQLRLYFDSSGNLLT

HAAdV-C5K	NSFLDPEYWNFRNGDLTEGTA YTNVAVGFMPNLSAYPKSHCK-----TAKSNIVSQVYL---NGDKTKPVTLLTIT
HAAdV-D26K	SSSLDKDYWNYSRDDSI VVSQKYNNAVPFMPNLTAYPKPSAQN-----AKNYSRTKIISNVYL---GALTYQPVIITIA
HAAdV-D48K	DSLLDKDYWNYSRNNNVGSA YEEAVGFMPSTTAYPKPPTPTPTNPTPLKKSQAKNKYVSNVYL---GGQAGNVPVATTS
HAAdV-D49K	GSLLDGNVWNYKNKDSVIGSPYENAVPFMPENSTAYPKIINNGTAN-PEDKKSAAKKTIVENVYL---GGDAAKPVATTS
HAAdV-B11K	RLSSLLKTPLNHKSGQNMATGAI TNAGFMPSTTAYPFDNSRE-----KENYIYGTCTCYTA-SGHTAFPIDISVM
HAAdV-B35K	EESDLKIPLKNKSSSTAT-SETVASSKAFMPSTTAYPFTTTRD-----SENYIHGICYMYSYDLSLFPNLISIM

HAAdV-C5K	LNGTQETGTPSAYSMSFSWDWSGHNYINE-----IFATSSYTFYSIAQE
HAAdV-D26K	FNQEEENG---CAYSITFTFWQ-KDYSAQ---QFVTSFTFSYLTQENKDKD
HAAdV-D48K	FNKE--TG---CTYSITFDFAWN-KTYENV---QFDSSFLTFYSIAQE
HAAdV-D49K	FNKEEESN---CVYSITFDFAWN-KTYKNV---PFDSSSLTFYSIAQDAEDKN
HAAdV-B11K	LNRR--T---ETSYCIRITWSWNTCDAPVQTSATTLVTSPPFTFYIREDD
HAAdV-B35K	LNSRM---VAYAIQFEWNLNASESPESN---IATLTSPPFFYSITEDDN

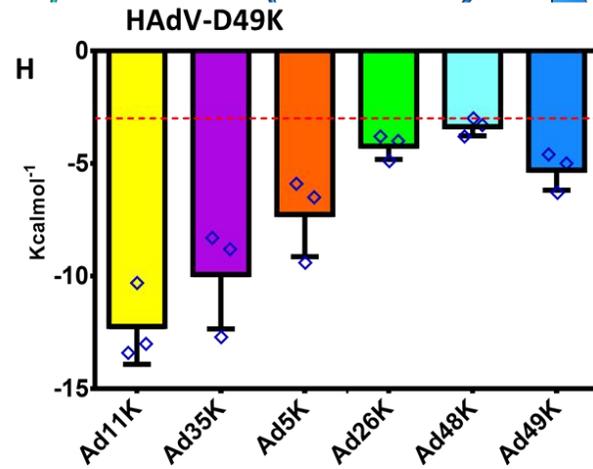


Figure 4-7 Species D adenovirus fiber-knobs are unlikely to form a stable complex with CD46. Homology modelling of adenoviruses in complex with CD46, based on the known crystal structure of HAdV-B11K in complex with CD46, shows interfaces at the HI, DG, GH, and IJ (From left to right, respectively) loops. HAdV-C5K, a non-CD46 utilising adenovirus species, is seen to form very few polar contacts (interacting residues shown as sticks, connected by red dashed lines) with CD46 when modelled (A). HAdV-B11K (B) and HAdV-B35K (C), both known CD46 utilising viruses, show large numbers of polar interactions with CD46. Similar modelling of HAdV-D26K (D), D48K (E), and D49K (F) reveals few polar contacts with CD46, similar to HAdV-C5K. Sequence alignment of the modelled adenovirus fiber knob proteins reveals significant sequence divergence in which the species D fiber-knobs share few, if any, residues associated with CD46 binding (highlighted in green) in HAdV-B11K or B35K (G). QT-PISA predictions of interface energy for each fiber-knob with CD46 concur, showing strong binding energies for the known CD46 binding knobs HAdV-B11k and 35K, weaker binding for the non-CD46 interacting HAdV-D5K, and very weak interaction for the species D fiber-knobs (H). Values are the mean of $n=3$ calculations based upon the 3 interfaces found in each native trimer, error is standard deviation from the mean. Dashed red line denotes $-3.0 \text{ Kcalmol}^{-1}$.

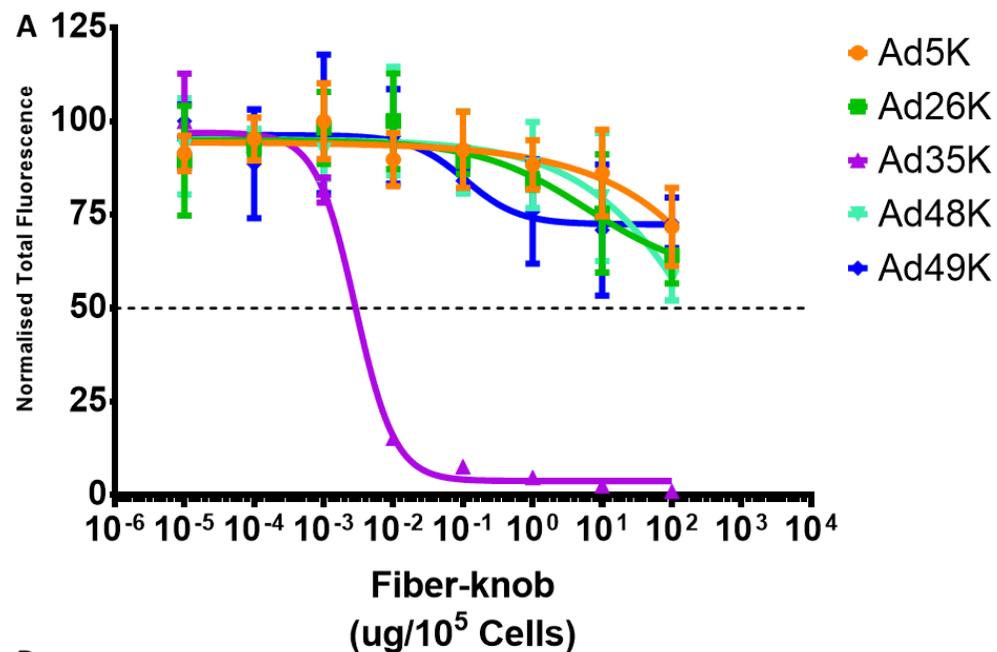
diverse regions of the protein, both at a 3-dimensional and a sequence level. Indeed, the species D adenoviruses do not share any of the same residues predicted to bind to CD46 as are in the species B controls.

Calculations of the binding energy between each of these fiber knob proteins and CD46 show that very low interface energies (indicating a more stable interaction interface) are found with HAdV-B11K and HAdV-B35K (Fig4-7E). HAdV-C5K exhibited a stronger interface than expected, which is likely due to the proximity of the relatively large HAdV-C5K fiber knob HI loop compared to the other proteins. The species D fiber knobs are all higher energy than that of the species B's, indicating a weaker interaction. HAdV-D26K and HAdV-D48K energies approach $-3.0 \text{ kcalmol}^{-1}$, which is the value expected for random proteins passing in solution. HAdV-D49K interaction with CD46 is similarly weak, but appears slightly stronger than the other two species D viruses, again likely due to the slightly more extended HI loop in which the contacts are seen (Fig4-7F,G,H).

Based on these computational simulations it appeared unlikely that any of the investigated species D adenoviruses could interact with CD46 in a stable manner. We sought to confirm this finding by conducting competition inhibition assays and measuring the IC_{50} concentration (the concentration of protein capable of inhibiting a response by 50%), using antibody binding as a read out.

The HAdV-D35K positive control strongly inhibited detection of CD46 using the corresponding antibody, with an IC_{50} of $300 \text{ ng}/10^5$ cells (Fig4-8A,B). The other tested fiber knob proteins were all incapable of reducing antibody binding

(measured as total fluorescent intensity) by 50% and followed a similar profile as HAdV-C5K negative control (Fig4-8A).



B

Fiber-Knob Species	Ad5	Ad26	Ad35	Ad48	Ad49
CD46 IC ₅₀ (μg/10 ⁵ Cells)	*nm	*nm	0.003	*nm	*nm

C

	CD46		
	K_{on} (M ⁻¹ s ⁻¹)	K_{off} (s ⁻¹)	K_D (μM)
HAdV-B3K	1.6x10 ⁵	0.025	0.16
HAdV-C5K	nb	nb	No binding
HAdV-B35K	1.2x10 ⁶	0.032	0.027
HAdV-D26K	nm	nm	50.8
HAdV-D48K	nb	nb	No binding
HAdV-D49K	nb	nb	No binding

nm = kinetics too fast to measure
nb = no binding

Figure 4-8 Competition inhibition assays suggest species D adenovirus fiber-knobs cannot bind CD46. IC₅₀ curves show potent inhibition of CD46 binding by anti-CD46 antibody by HAdV-B35K, whilst the non-CD46 utilising HAdV-C5K follows a similar profile to the species D fiber-knob proteins which are all incapable of effectively inhibiting CD46 binding by the antibody (A). HAdV-B35K has an IC₅₀ value of 0.003 μg/10⁵ cells (B), no other fiber-knobs tested were capable of achieving 50% inhibition. SPR measurements were performed using recombinant virus fiber knob and CAR, showing specific interactions between all the tested species D and C fiber-knob proteins (C). n=3 biological replicates, curve fitted by variable slope least squares regression. Short form names used for readability (Ad5 is HAdV-C5, etc).

We used surface plasmon resonance to make a direct measurement of any affinity between the fiber knobs and CD46. As expected HAdV-D35K bound tightly ($K_D = 0.027\mu\text{M}$), while HAdV-B3K bound CD46 with a lower affinity ($K_D = 0.16\mu\text{M}$). This is consistent with previous reports regarding the two viruses. Surprisingly HAdV-D26K was capable of CD46 interaction, albeit with weak affinity ($K_D = 50.8\mu\text{M}$), more than 1800x lower affinity that of HAdV-B35K. however, the precise kinetics of the HAdV-D26K interaction were too fast to measure, suggesting a highly unstable interaction (Fig4-8C).

Table 4-2 Surface plasmon resonance (SPR) measurements confirm finding regarding CD46 and Car affinity and rule out desmoglein 2 as a potential receptor for these species D adenoviruses. The on (K_{on}) and off (K_{off}) rate of the indicated adenovirus fiber-knob proteins were measured for CAR, CD46, and DSG2 and used to calculate the affinity (K_D).

	DSG2		
	$K_{on} (M^{-1} s^{-1})$	$K_{off} (s^{-1})$	$K_D (\mu M)$
HAdV-B3K	nm	nm	66.9
HAdV-C5K	nb	nb	No binding
HAdV-B35K	nb	nb	No binding
HAdV-D26K	nb	nb	No binding
HAdV-D48K	nb	nb	No binding
HAdV-D49K	nb	nb	No binding

nm = kinetics too fast to measure
 nb = no binding

4.3.2 Desmoglein 2 is unlikely to be the primary receptor for the investigated species D adenovirus fiber knobs:

We also investigated the ability of the HAdV-D26/48/49K to bind to DSG2. Surface plasmon resonance showed interaction between HAdV-B3K, the best studied of the DSG2 binding viruses (Table4-2). While the interaction was not highly stable, as demonstrated by the fast kinetics, we were able to use the

single injection method at multiple concentrations to determine the affinity. The same experiment with HAdV-D26K, HAdV-D48K, and HAdV-D49K showed no interaction in all cases, seemingly ruling out DSG2 as a potential viral receptor (Table4-2).

4.3.3 Adenoviruses D26, D48, and D49 fiber knob proteins bind CAR with low affinity and suggest regulation through steric hindrance:

Modelling of the fiber knob proteins in complex with CAR was performed in a similar way as with CD46, using the previously published structure of HAdV-D37K in complex with CAR as a template (PDB 2J12)⁷³. The energy minimised models of fiber knobs in complex with CAR (Fig4-9A) show the conservation of the fiber knob core fold and the two major interfaces (α and β) between CAR and the virus proteins.

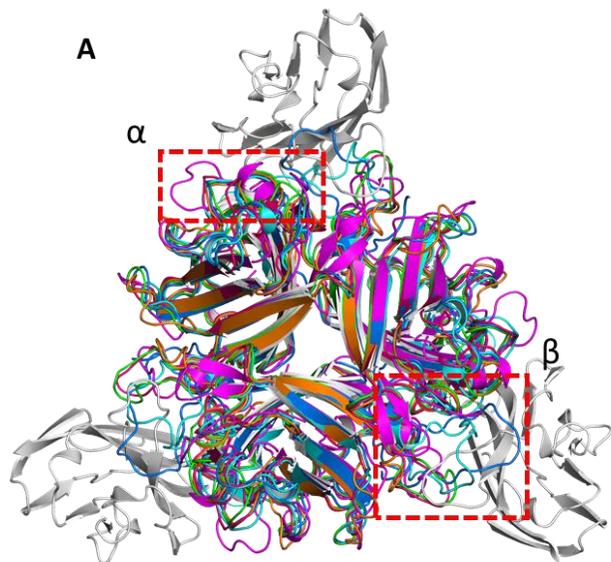
Amino acid sequence alignment shows that the different fiber knob proteins all form numerous polar contacts to CAR throughout the α -interface, while the non-CAR binding HAdV-B35K negative control shows few (Fig4-9B). Residues which have previously been shown to be important for CAR interaction (Kirby *et al*⁷⁵) are seen to be partially conserved. Similarly, there is a high degree of similarity between polar contact forming residues for the various species C and D fiber knobs at both a sequence (Fig4-9B) and tertiary structure level (Fig4-9C).

Calculation of the interface energy between CAR and the fiber knob α -interface showed all the species D adenovirus fiber knobs had more stable binding interfaces than the non-CAR binding HAdV-B35K fiber knob. This data regarding the α -interface therefore suggests that HAdV-D26/48/49K should be capable of interacting with CAR.

However, the β -interface shows a significant clash between the DG-loops of HAdV-D26/48/49K and the apical facing domain of CAR (Fig4-9E). HAdV-C5K, the prototypical CAR binding adenovirus, and HAdV-D37K do not form a clash at the β -interface. However, HAdV-D26K forms a small clash, while the more extended DG loops of D48K and D49K form more significant clashes.

Based upon this clash we would expect to see a reduced affinity for CAR in species D adenoviruses with longer loops due to a propensity for the loop to

prevent stable interface formation when extended in a conformation which would clash with CAR. This was confirmed through competition inhibition, SPR, and time course infectivity experiments.



B

HAdV-C5K	400	TLWTTPEPSENCRLNAE---KDAKLTLLVLTKCGSQILATVSVLAVKGSGLAPISGTVQ-----SAHLIIRFDENGVLLNN
HAdV-B35K	133	TLWTTGIPPPNCQIVENTNTDGLKLTLLVLVKNGLLVNGVSVLVGVSDTVNQMFQTK-----TANIQLRLYFDSSGNLLTE
HAdV-D37K	184	TLWTTTPTDTPNCTIAQD---KDSKLTLLVLTKCGSQILANVSLIVVAGKYHIIINNKTNPK---IKSFTIKLLFNKNGVLLDN
HAdV-D26K	188	TLWTTTPTDTPNCKMSTE---KDSKLTLLVLTKCGSQVLGNVSLAVTGEYHQMTATT-----KKDVKISLLFDENGILLPS
HAdV-D48K	182	TLWTTTDDPSPNCKIDQD---KDSKLTFLVLTKCGSQILANMSLLVVKGKFSMINNKVNGTDDYKKFITIKLLFDEKGVLLKD
HAdV-D49K	188	TLWTTTDDPSPNCKVSEE---KDSKLTLLVLTKCGSQILASVLLVVKGKFANINNKTNPGEDYKKF SVKLLFDANGKLLTG

HAdV-C5K	SFLDPEWVNFNRGDLTEGTAYTNAVGFMPNLSAYPKSHG-----KTAKSNIVSQVYLN---GDKTKPVTLTITL
HAdV-B35K	ESDLKIPLKNKSS-TATSETVASSKAFMPSTTAYPFNTTRDS-----ENY---IHGICYMTYDRSLFPLNISIML
HAdV-D37K	SNLGKAYWVNRSGNSNVSTAYEKAIGFMPNLVAYPKPSNS-----KKYARDIVYGTIYLG---GKPDQPAVIKTFE
HAdV-D26K	SSLSKDWVNYRSDSIVSQYNNNAVFPMPNLTAYPKPSAQN-----AKNYSRTKIISNVYLG---ALTYQPVIIITIAF
HAdV-D48K	SSLDKEVWNYRSNNNNVGSAYEEAVGFMPSTTAYPKPTPTPTNPTTPELEKSQAQKNKYVSNVYLG---GQAGNEVATTVSF
HAdV-D49K	SSLDGNWVNYKNKDSVIGSPYENAVFPMPNLTAYPKIINNGTAN-PEDKKSAAKKTIVTNVYLG---GDAAKPVATTISF

HAdV-C5K	NGTQETGDTTPSAYSMSFSWDSGHNYI---NEIFATSSYTFYSYIAQE----
HAdV-B35K	NSRMIS---SNVAYAIQFEWNLNASESPESIATLTTPSPFFFSYITEDDN---
HAdV-D37K	NQE--TG---CEYSITFNFSWS-KTYE---NVEFETTSFTFSYIAQE----
HAdV-D26K	NQETENG---CAYSITFTFTWQ-KDYS---AQQFDVTSFTFSYLTQENKDKD
HAdV-D48K	NKE--TG---CTYSITFDFAWN-KTYE---NVFDSSELTFYSYIAQE----
HAdV-D49K	NKETESN---CVYSITFDFAWN-KTYK---NVPFDSSELTFYSYIAQDAEDKN

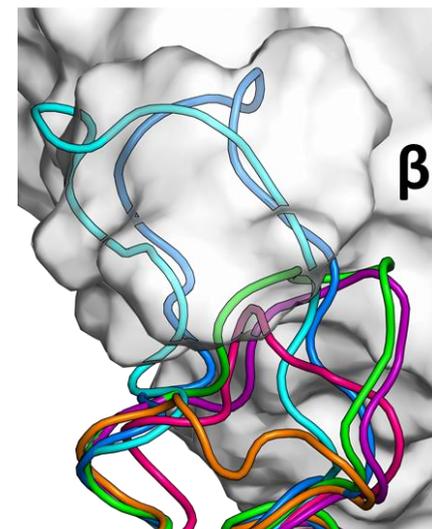
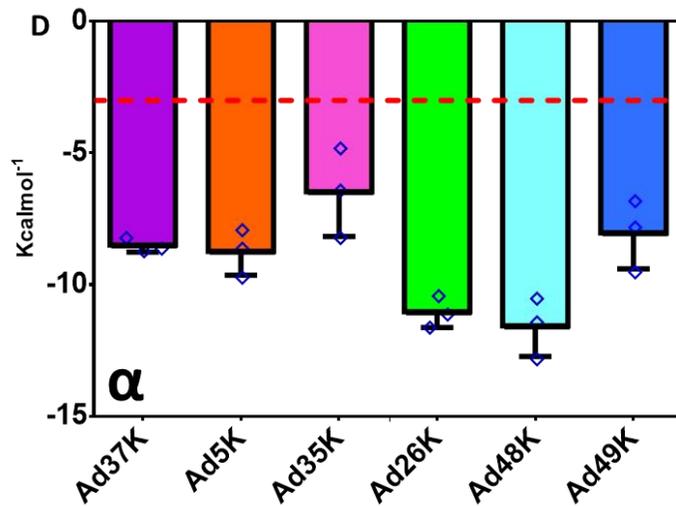
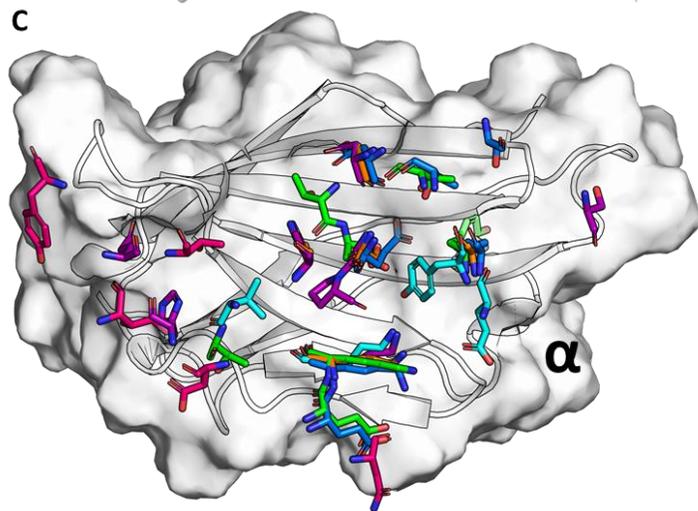


Figure 4-9 Adenovirus fiber-knob interaction with coxsackie and adenovirus receptor (CAR) is made up of two interfaces, and steric hinderance by the DG loop of species D adenoviruses may inhibit CAR interaction. Simulation of adenovirus fiber-knobs (coloured cartoon representations) in complex with CAR, based upon the existing structure of HAdV-D37K in complex with CAR (Grey cartoon), revealed two interfaces (**A**): α and β (red boxes).

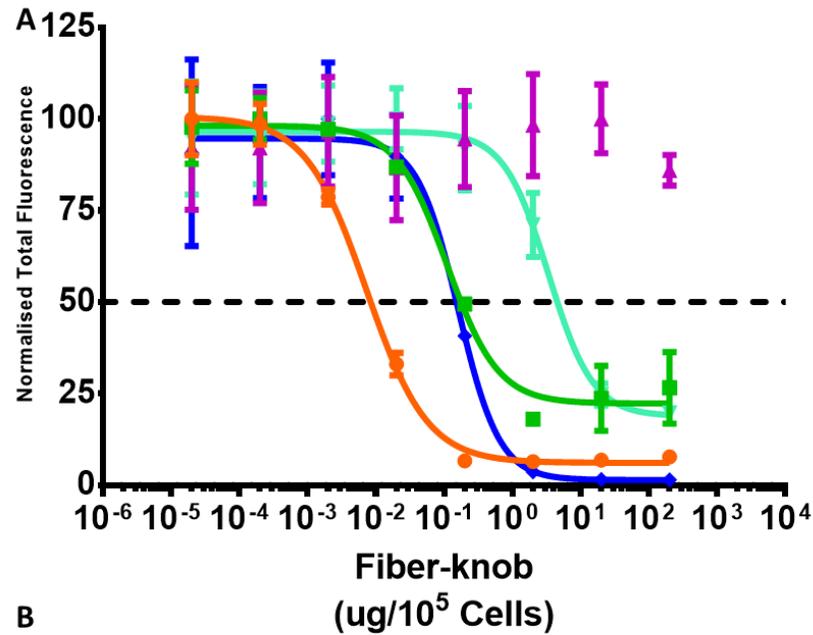
Fiber-knob residues which interact with CAR (highlighted in blue) show a high degree of conservation at both an amino acid sequence. Residues previously found to be critical for CAR interaction in HAdV-C5K (Kirby et al) are underlined. (**B**, Clustal Omega alignment of fiber-knob amino acid sequences) and a structural level (**C**, Grey surface is the maximum spatial occupancy of CAR, sticks are predicted CAR interaction residues coloured by virus). QT-PISA calculations of the interaction energy between CAR and the fiber-knob α -interface suggest stable interaction for all investigated fiber-knobs (**D**), however inspection of the β -interface shows severe clashes between CAR and the DG-loop of the fiber-knob (**E**). Values are the mean of $n=3$ calculations based upon the 3 interfaces found in each native trimer, error is standard deviation from the mean. Dashed red line denotes $-3.0 \text{ Kcalmol}^{-1}$.

Competition inhibition assays show that HAdV-C5K has the strongest affinity for CAR ($IC_{50} = 0.007\mu\text{g}/10^5$ cells). HAdV-D26K and D49K have approximately 100x lower affinities, with IC_{50} values of $0.110\mu\text{g}/10^5$ cells and $0.160\mu\text{g}/10^5$ cells, respectively (Fig4-10A,B). The weakest IC_{50} was observed for HAdV-D48K at $3.359 \mu\text{g}/10^5$ cells (approximately 500x weaker than HAdV-C5K), while the non-CAR binding HAdV-B35K control did not produce an IC_{50} value. SPR experiments confirmed the weak affinities of HAdV-D26/48/49K, however these experiments showed that it was HAdV-D49K had a stronger of the three affinities to CAR, with HAdV-D26K and HAdV-D48K appearing similar (Fig4-10C).

This contradiction cannot be explained by differences in the β -interface loop lengths alone, leaving methods of determination we used as a possible explanation. In the IC_{50} experiments the level of inhibition is a function of the affinity of the antibody used to infer blocking and the affinity of the recombinant fiber knob. In SPR the measurement is based upon an artificial environment, and affinities can be affected by myriad factors including flow rate and receptor density on the sensor chip.

In order to clarify the biological significance of the different affinities for CAR we performed an infectivity time course in which CHO-CAR cells were exposed to virus for different lengths of time. CHO-CAR cells express no known adenovirus receptor except for CAR, making them an important test cell line for HAdV:CAR interaction. Replication deficient HAdV-C5K pseudotypes were used to control for replication rate by preventing duplication of the transgene, and other non fiber knob mediated effects. If viral infection via CAR is the function of a

successful collision which is hindered by the DG loop at the β -interface then the
viruses with



B

Fiber-Knob Species	Ad5	Ad26	Ad35	Ad48	Ad49
CAR IC ₅₀ (μg/10 ⁵ Cells)	0.007	0.110	*nm	3.359	0.160

*not measured

C

	CAR		
	K_{on} ($M^{-1} s^{-1}$)	K_{off} (s^{-1})	K_D (μM)
HAdV-B3K	nb	nb	No binding
HAdV-C5K	2×10^6	0.0016	0.00076
HAdV-B35K	nb	nb	No binding
HAdV-D26K	3.4×10^2	0.0074	21
HAdV-D48K	2.5×10^2	0.0024	9.3
HAdV-D49K	6.1×10^3	0.0012	0.19

nm = kinetics too fast to measure

nb = no binding

Figure 4-10 Competition inhibition assays suggest species D adenovirus fiber-knobs bind to CAR with variable affinity. IC_{50} curves show potent inhibition of CAR binding by anti-CAR antibody by HAdV-C5K, whilst the non-CAR utilising HAdV-B35K is incapable of effectively inhibiting CAR binding by the antibody. The species D fiber-knobs are all capable of inhibiting CAR binding by the antibody but require greater amounts of protein than HAdV-C5K suggesting a weaker affinity (A). HAdV-C5K has an IC_{50} value of $0.007 \mu g/10^5$ cells, the species D fiber-knobs were between 2 and 3 orders of magnitude less inhibitory (B). SPR measurements were performed using recombinant virus fiber knob and CAR, showing specific interactions between all the tested species D and C fiber-knob proteins (C). $n=3$ biological replicates, curve fitted by variable slope least squares regression. Short form names used for readability (Ad5 is HAdV-C5, etc).

longer loops should require longer times to infect, which is what was observed in the assay (Fig4-11).

HAdV-C5K infected CHO-CAR cells rapidly, infecting 80% of cells with only an initial exposure (0mins, on ice) and rapidly approaching 100% infectivity (Fig-12A), which the non-CAR utilising HAdV-C5/B35K was unable to achieve efficient transduction (Fig4-11B). HAdV-C5/D26K, possessing the shortest loop of the 3 species D fiber knobs of interest, was the next most efficient with an initial infectivity of 65% (Fig4-11C). HAdV-C5/D48K and HAdV-C5/D49k, possessing the longest loops, were the least efficient requiring between 10 and 30 minutes to improve over the initial infectivity. HAdV-C5/D48K required 180mins to reach maximum infectivity (Fig4-11D). HAdV-C5/D49K did not reach 100% infection, an affect which has been observed repeatedly in CHO-CAR cells and is yet to be explained. However, the HAdV-C5/D49K seems to follow the same rate of infection profile as HAdV-C5/D48K, which as a similarly long loop.

Taken together these results support the hypothesis that the length of the DG loop directly impacts CAR interaction with the adenovirus Fiber knob by steric hindrance at the β -interface. This leaves in question why HAdV-D48K is seen to have a 21x weaker inhibitory effect than HAdV-D49K (Fig4-10A), despite theoretically creating less steric hindrance due to the HAdV-D48K DG loop being 3 residues shorter.

The answer to this question may be due to the composition of the HAdV-D48K DG-loop (Fig4-12A). The HAdV-D48K loop contains numerous polar bonds which help to restrain its kinked, hairpin like, tertiary structure (Fig4-12A,B). The N-terminal half of this loop has many proline residues which provide increased rigidity compared to other residues. Combined with 5 polar bonds to the opposing fiber knob monomer with the non-proline rich half of the loop, this may create a situation which holds the HAdV-D48K DG loop in a conformation inhibitory to CAR binding, reducing the affinity further than it would were it more flexible.

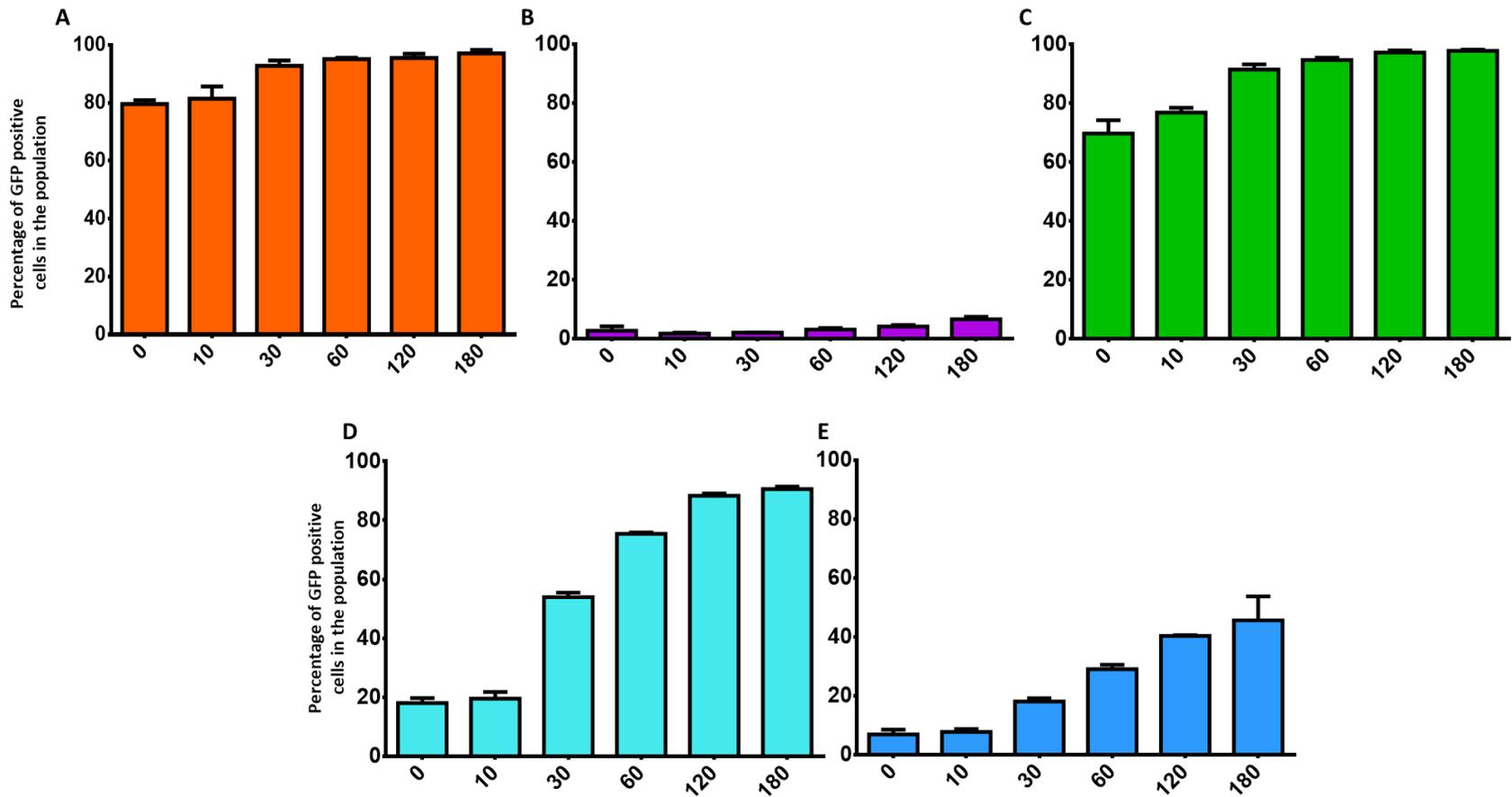


Figure 4-11 Efficiency of infection by adenovirus fiber-knob pseudotypes is proportional to incubation time in CHO-CAR cells. *HAdV-C5* is a known CAR utilising adenovirus and is seen to infect CHO-CAR cells efficiently immediately upon contact with further incubation slightly improving the percentage of infected cells (A), whereas *HAdV-C5/B35K* (*HAdV-C5* pseudotyped with the fiber-knob domain of *HAdV-B35*) which binds CD46 was unable to efficiently infect even with a long incubation (B). *HAdV-C5/D26K* was more affected by short incubation times (C), but not as significantly as *HAdV-C5/D48K* which did not maximally transform cells until two hours of incubation (D). *HAdV-C5/D49K* was most affected, however did not achieve complete transduction of the cells at any time point, despite appearing to have neared maximum transformation efficiency between 2 and 3 hours of incubation (E). *n*=3 biological replicates.

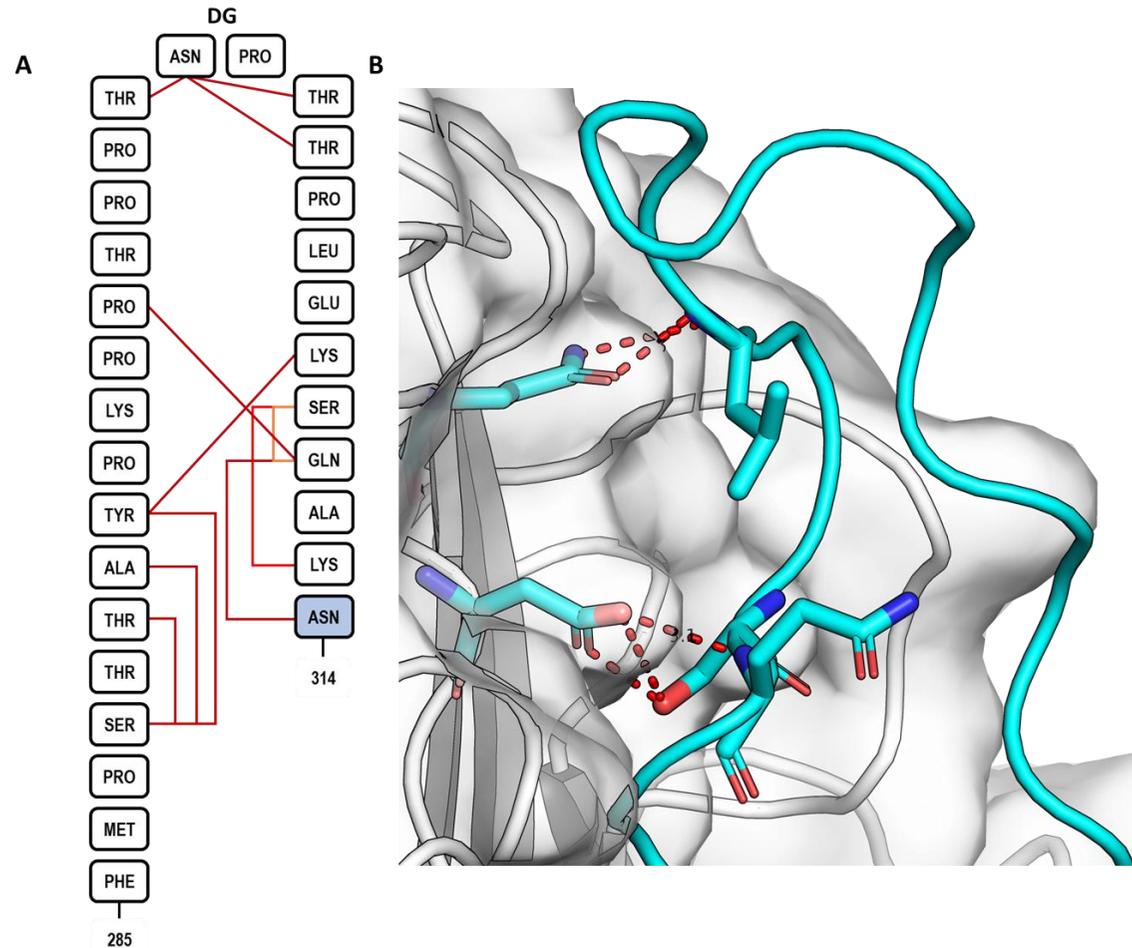


Figure 4-12 The proline rich loop of HAdV-D48K contributes to a low CAR affinity. The HAdV-D48K DG-loop is highly proline rich and forms polar contacts (red lines are single bonds, orange lines are two bonds) between residues which help retain its' tertiary structure (**A**). The DG loop (cyan cartoon) forms further polar contacts to residues in the opposing monomer (contact forming residues shown as sticks, opposing monomer in white) which stabilise its location in the inter-monomer cleft (**B**). Red atoms are oxygen, blue atoms are nitrogen, structure is PDB 6FJQ.

4.4 Discussion:

The data in this study demonstrates that none of the three investigated species D adenovirus proteins, HAdV-D26K, HAdV-D48K, HAdV-D49K, are unlikely to utilise CD46 or DSG2 as a primary receptor. Instead we show that the viruses are capable of low affinity CAR engagement. This is in direct contradiction of previous studies suggesting CD46 as the primary receptor for HAdV-D26 and HAdV-D49K, this has implications for the therapeutic use of vectors based upon these viruses.

We demonstrated the ability to purify recombinant fiber knob protein to high purity and generate multimeric protein of the correct size and multimer number (3-mer) to be a fiber knob protein (Fig4-1). In both chromatograms minor peaks are visible at the start (Fig4-1A) or end (Fig4-1C). The low molecular weight bands visible on the gels are also intriguing, as they appear despite SEC, which should remove proteins of significantly different molecular weights. A possible explanation for both the minor peaks and low molecular weight bands is that there is breakage occurring in the proteins as they are dragged through the matrix, either of the gel or column. It seems probably that these are the tail domain of protein containing a 6-HIS purification tag and 13 residues of the fiber knob shaft domain. This tail is left exposed, protruding from the globular structure of the fiber knob trimer and is thus likely to be more prone to breakage. This would explain the minor peak at the beginning of the Ni-IMAC chromatogram, as small his tagged peptides would elute easily compared to relatively large proteins and would have a low absorbance due to their small size, which explain the presence of a small peak at the end of the SEC chromatogram as the small peptides are retained on the column longer. It also offers an explanation for the low molecular weight bands presence even after SEC, as further tails could be cleaved from the fiber knob during the running of the gel.

We describe the crystallisation of multiple fiber knob proteins (Table.1), including two previously undescribed structures (Fig4-2,3) and a higher resolution version of HAdV-C5K (Fig4-4). The three different species of fiber knob protein described in this study crystallised in three different space groups. This might be surprising given the similar overall morphology of the proteins,

with the most significant difference between them being the size of the loops. These loops are likely to be what is disrupting the packing arrangement as we see the proteins with larger loops (HAdV-D49K and HAdV-D48K) have lower level symmetry compared to the more compact fiber knobs such as HAdV-D26K. The interesting feature, in terms of biological impact, is that this suggests the loops are not fully flexible, and instead have preferred conformations to an extent which dictates the crystal packing arrangement. This implies that the loops are not the fully flexible motifs they have sometimes been purported to be, but instead are evolved motifs with a distinct structure. This is further supported by the tertiary structure observed in the DG-loop of HAdV-D48K (Fig4-12A,B).

Despite the differences in the crystallographic arrangement, we could solve the structures of all 4 of these species. This provides new insight into the differences between the primary tropism determining proteins of the understudied species D viruses. We calculated that different virus serotypes fiber knob domains have different interface energies (Fig4-6). Since we have not determined the interface energies of the fiber-shaft domains we cannot determine the overall strength of the trimeric interface of the fiber proteins. However, this finding shows that trimeric interface stability is not consistent throughout the adenoviruses, the implications for this variable level of trimer stability for the virus' life cycle are unclear.

Previous studies have asserted that CD46 is the primary receptor for HAdV-D26¹⁰³, HAdV-D48³⁷, and HAdV-D49³⁵¹. Our computational simulations based on the crystal structure of the primary receptor determining fiber knob protein of these three viruses suggested that it was unlikely that they would be able to form a stable interface (Fig4-9), and without a stable interface the virus cannot attach to the cell to initiate infection. More factors than just polar contacts can be important in interface formation, including hydrophobic interactions and electrostatic interactions as seen previously³⁷⁴, though they are unlikely to be a critical determinant in such a solvent exposed loop. To account for these factors, we performed interface energy calculation with QT-PISA to determine the total binding interface energy which confirmed a weak interface with CD46 for the three species D adenovirus fiber knobs (Fig4-9H).

An unexpected observation in the QT-PISA calculations was a surprisingly strong predicted interface energy between CD46 and HAdV-C5K (Fig4-9H). HAdV-C5K has a long HI-loop compared to the other investigated fiber knobs. This increase length means that, after energy minimisation, the loop invades between CD46 motifs creating a large surface to surface contact area. This results in QT-PISA measuring an artificially large interaction interface which is unlikely to exist in nature due to the energetic requirements of creating such a large shift in the tertiary structure of CD46 to enable the HI loop penetration. This action by the HAdV-C5K HI-loop has a further effect of creating a large distance between the other loops and CD46 precluding any interaction with them (Fig4-9C).

The relative position of the four critical loops (HI, DG, GH, IJ) is important in successful CD46 interactions, not just in the prevention of the HAdV-C5K interaction. CD46 interaction between HAdV-B11K has been shown to be two orders of magnitude stronger than the interaction HAdV-B16K due to a 2 amino acid insertion in the FG-loop of HAdV-B16K requiring a greater angle between the SC1 and SC2 domains of CD46 to enable binding³⁷¹. A similar mechanism was reported for HAdV-B21K, where a shorter IJ loop was also seen to be important (as was theorised in the HAdV-B16K study) by enabling closer interaction⁵¹. The adenoviruses with shorter FG-loops (synonymous with the DG-loop in other adenoviruses) had higher affinities than those with longer loops. HAdV-B35K (KD = 19nM) and HAdV-B11K had dissociation constants of 19nM and 13nM, respectively, compared to the long looped HAdV-B16K and HAdV-B21K which had dissociation constants of 437nM and 281nM.

The DG-loops of HAdV-D26/48/49K are longer than the DG(FG)-loops of HAdV-B11K and HAdV-B35K by 3, 10, and 9 amino acids respectively. Since just a 2 amino acid increase in length is enough to reduce the CD46 binding capability of the species B adenoviruses by >10x it seems likely that the limiting effect of the CD46 SC1-2 bond angle would preclude an effective interaction with these species D adenoviruses even if they did share crucial CD46 interacting residue identities, which they do not (Fig4-9G).

This hypothesis was confirmed by biological assay, in which no quantity of recombinant fiber knob protein was capable of completely inhibiting α CD46

antibody binding up to $100\mu\text{g}/10^5$ cells (Fig4-10A,B). Interestingly, while competition inhibition assay showed no binding, SPR showed that HAdV-D26K had low affinity for CD46 ($K_D = 50.8\mu\text{M}$) (Fig4-10C). This unexpected finding places the CD46 binding affinity of HAdV-D26K $>10x$ lower than the weakest of the measured species B adenoviruses, HAdV-B16K³⁷¹. Our own SPR analysis of CD46 binding places HAdV-D26K affinity for CD46 $>300x$ weaker than that of HAdV-B3K, another known weaker CD46 binding adenovirus (Fig4-8C)³⁷⁵. This suggests that plasticity is allowed in CD46 interaction, with loop binding be a more significant factor than would previously have been thought. It seems unlikely that HAdV-D26K can utilise CD46 as a primary receptor tropism when the affinity is so low and it is incapable of stable interaction under biological conditions (Fig4-10A,B).

HAdV-B3K, the best studied of the DSG2 binding adenoviruses, showed low affinity interaction with DSG2. Low affinity is to be expected since it has previously been observed that HAdV-B3 requires polyvalence suggesting an avidity effect is required for the functional utilisation of DSG2 as a receptor^{54,55,110,125}. Regardless of the need for polyvalence, there was a measurable direct interaction between HAdV-B3K and DSG2. None of the tested species D adenovirus knobs were found to be capable of interacting with DSG2, making its use as a primary receptor for these viruses highly improbable.

Modelling of the complexes with CAR showed a two-part interface. At one end of the protomer a large relatively flat region of the fiber knob, termed the α -interface, is seen to be able to form a CAR binding interface (Fig4-11A,C,D). This supports previous observations of plasticity within the CAR binding interface⁷³.

On the opposite end of the protomer, but immediately opposite the α -interface due to the quaternary protein structure, is the β -interface. This β -interface is seen to act to reduce CAR binding affinity through steric hindrance (Fig4-11E), protruding from the intermonomer cleft and, presumably, preventing collision in an orientation which enables stable complex formation. This hypothesis is further supported by the interaction data showing that, in a biological context, viruses with DG-loops (the critical determinant at the β -interface) with greater occupancy of the intermonomer cleft infect cells less efficiently (Fig4-10-14).

Taken together, these data suggest that the species D adenoviruses are capable of engaging CAR as a receptor. HAdV-D26K is only $\approx 15\%$ less infectious than HAdV-C5K in the time course infectivity assay (Fig4-13A,C). HAdV-C5/49K is seen to infect CHO-CAR cells with a similar profile to that of HAdV-D48K, which is unsurprising given the similar DG-loop lengths (Fig4-11D,E). What is more surprising is the inability of HAdV-C5/49K to infect CHO-CAR cells beyond 50%. The reason for this seemingly arbitrary lack of increased infectivity over time is unclear. However it has previously been observed in an assay utilising the whole HAdV-D49 serotype that when CAR is present HAdV-D49 has a lower infectivity³⁷. The reason for this is unclear, however we know that HAdV-C5/49K does not require CAR to infect cells (See later), so this oddity of infection kinetics may be a result of an alternative mechanism of infection mediated by an as yet unidentified receptor.

Despite these viruses displaying the ability to infect via CAR, their engagement of CAR as a primary receptor makes little sense from an evolutionary perspective. Recombination events are common among adenoviruses^{20,21,32,376}, so isolated instances of viruses with sub-optimal affinities for their primary receptors are to be expected. However, HAdV-D26K is persistent in the Japanese population, being a cause of EKC³⁷⁷. While HAdV-D49 does not require CAR for infection³⁵¹, and data shows CAR actually diminishes HAdV-D49³⁷ and HAdV-C5/49K infections (see chapter 3). Therefore, given the evolutionary pressure to maximise infectivity and subsequent virus propagation, it is counterintuitive that these viruses would persist in the population without mutating/recombining to eliminate these loops which hinder their infection. We therefore suggest that these adenoviruses must use an alternative molecule on the cell surface as their primary means of infection, alleviating the negative consequences of this CAR inhibitory loops.

An HAdV-D26 based virotherapy which is currently in phase-III clinical trials as a vaccine for Ebola¹⁰⁵, and as protective agents against HIV^{104,378} and Zika³⁷⁹ has shown excellent efficacy and potent immunogenicity. Similarly HAdV-D49K derived vectors are seen to be highly infectious, both in our own studies and those of other teams, making them effective gene transfer vehicles^{350,351,380}.

Previously it has been thought that HAdV-D26¹⁰³, HAdV-D48^{41,381}, and HAdV-D49³⁵¹, viruses utilise CD46, it now seems unlikely^{72,79,382}. The receptor which these viruses engage plays an important role in the mechanism of action of vectors derived from viruses. Elucidation and characterisation of these virus: receptor interactions, therefore, represents something of a priority, such that the phenotypes which make them useful vectors (immunogenicity, and efficient infectivity) might be exploited for therapeutic application.

4.5 Summary of Chapter 4

The major findings in this chapter have been:

1. Comparison of adenovirus fiber knob protein structures can predict CAR and CD46 binding ability.
 - a. Predicted by computational simulation (Fig4-9), then confirmed by competition inhibition assays and SPR analysis (Fig4-10)
2. None of the investigated species D adenoviruses appear to utilise CD46 as a primary cell entry receptor.
 - a. Shown by computational prediction of CD46 binding interfaces (Fig4-7), competition inhibition assays and SPR analysis (Fig4-8).
3. None of the investigated species D adenovirus fiber knob proteins appear to interact with desmoglein 2.
 - a. Confirmed by SPR analysis (Table 4-2).
4. All the tested species D adenoviruses appear to be capable of using CAR as a cell entry receptor, but with low affinity.
 - a. As confirmed by competition inhibition assays and SPR analysis (Fig4-10)
5. The DG loop likely influences CAR binding affinity by sterically hindering interaction at the β -interface.
 - a. Longer DG loops, or those with less flexibility to move out of a conformation which obstructs binding, will be more inhibitory to CAR binding (see discussion).
 - b. This is inferred by analysis of the time course infectivity assay in Figure 4-11, the preferred conformation of the DG loop of HAdV-D48K, and knowledge of the DG loop published previously⁷².

CHAPTER 5. Results: Evaluation of Species D adenovirus fiber knob tropisms

5.1 Introduction:

Sialic acid is a known adenovirus receptor for multiple species D types, the first description being in HAdV-D37/64 (The fiber knob of HAdV-D37 and HAdV-D64 are identical) and HAdV-D19p, which differs with HAdV-D37 fiber knob by just 2 amino acids^{56,146,383}. This was further explored by crystallography, which demonstrated sialic acid binds to the apical depression of these fiber knobs²⁷. However, apical binding is not the only mechanism by which adenovirus can bind to sialic acid, with turkey siadenovirus 3¹³² and canine adenovirus 2⁸² binding sialic acid in lateral regions. Lateral sialic acid binding is also reported in Human species G adenovirus type 52⁷¹, which has been shown to be especially dependent on polysialic acid⁵⁷, a rare glycan motif which is associated with neuronal adhesion proteins^{141,144} although may not be expressed there exclusively¹³⁶. HAdV-D37 also exhibits a glycan preference, utilising to GD1a to infect cells¹⁴⁷, although it seems unlikely to be a strictly specific interaction, as sulfated glycosaminoglycans (GAGs) having been shown to be effective decoy receptors, reducing viral infectivity^{364,365}.

Based upon the work with HAdV-D37, a known cause of epidemic keratoconjunctivitis, and infectivity assays performed in human conjunctival epithelial (HCE) cells³⁸⁴, sialic acid has been presumed to be the primary means of infection for EKC causing adenoviruses. The most severe varieties of adenovirus keratoconjunctivitis are caused by HAdV-D8^{321,385,386}, 37^{321,387,388}, and 64 (formerly HAdV-D19a)^{18,321}. However, recent outbreaks have been associated with HAdV-D53 (previously classified as HAdV-D22/H8)^{32,389}, HAdV-D54³⁹⁰, and HAdV-D56³⁹¹, with further instances reported for other species D adenovirus including HAdV-D26 (formerly the Kobe-H strain) and HAdV-D49 (formerly the Kobe-S strain)^{345,346,377}. Usage of sialic acid has been confirmed in these EKC causing adenoviruses in recent work⁵⁹.

After eliminating known primary adenovirus receptors as the primary means of infections for HAdV-D26 we explored the potential for the usage of sialic acid, based on structural homology to known sialic acid utilising EKC causing adenoviruses. In this study we demonstrate sialic acid as a primary receptor for

HAdV-D26K and explore the possibility of sialic acid and other sugars as a receptor in other species D adenoviruses.

5.2 HAdV-D26 fiber knob has homology with known sialic

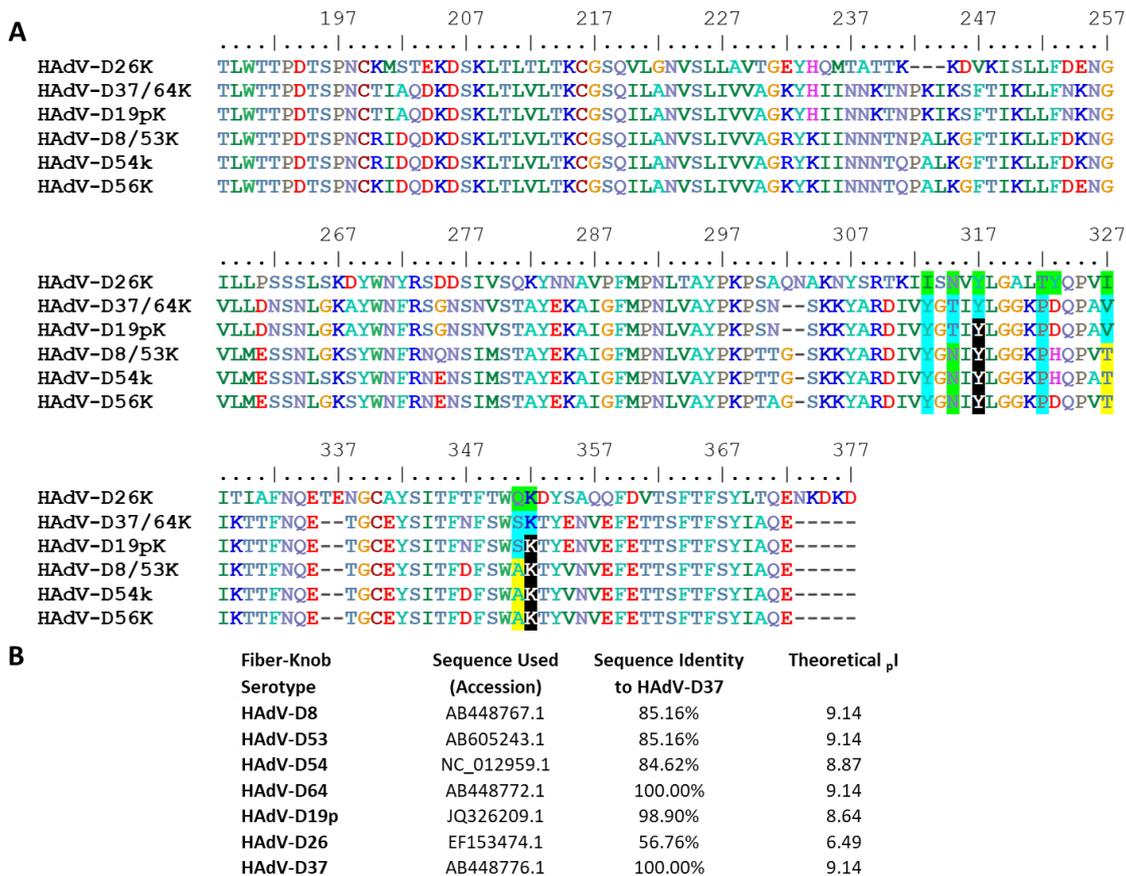


Figure 5-1 HAdV-D26K shares key sialic acid binding residues with EKC associated adenoviruses (A) but has low overall sequence identity and a very different predicted isoelectric point (B).

acid utilising adenoviruses

To explore potential receptor tropisms of HAdV-D26 we investigated the fiber knob protein, as it is the primary tropism determinant which forms initial interactions on the cell surface, in all known adenoviruses. Sequence alignment of HAdV-D26K shows homology to known sialic acid utilising adenoviruses (Fig5-1A). Crucially, the Lys352 and Tyr317 in HAdV-D26 are conserved. These residues have been shown to be critical to stable sialic acid binding in other adenoviruses²⁷. Other residues which are important for sialic acid binding in HAdV-D37 are not conserved in HAdV-D26K but have similar character. For example, Ile313 and Ile327 are not identical but are still hydrophobic like the tyrosine and valine at the equivalent, respective, positions in HAdV-D37. Interestingly Asn315 is not conserved between HAdV-D37 and HAdV-D26, but the asparagine is the more common residue in the EKC causing viruses despite

HAdV-D37 being the classical and best studied of the sialic acid utilising adenoviruses (Fig5-1A).

However, whole sequence level analysis of HAdV-D26K in comparison to the other EKC causing viruses highlights greater differences. The EKC viruses all have sequence identity to HAdV-D37K of >84%, whilst HAdV-D26K is <57%, showing the level of divergence between HAdV-D26K and the EKC causing viruses fiber knobs at the primary sequence level (Fig5-1B). Similarly, all the EKC viruses have basic predicted isoelectric points of >pH8.6, whilst HAdV-

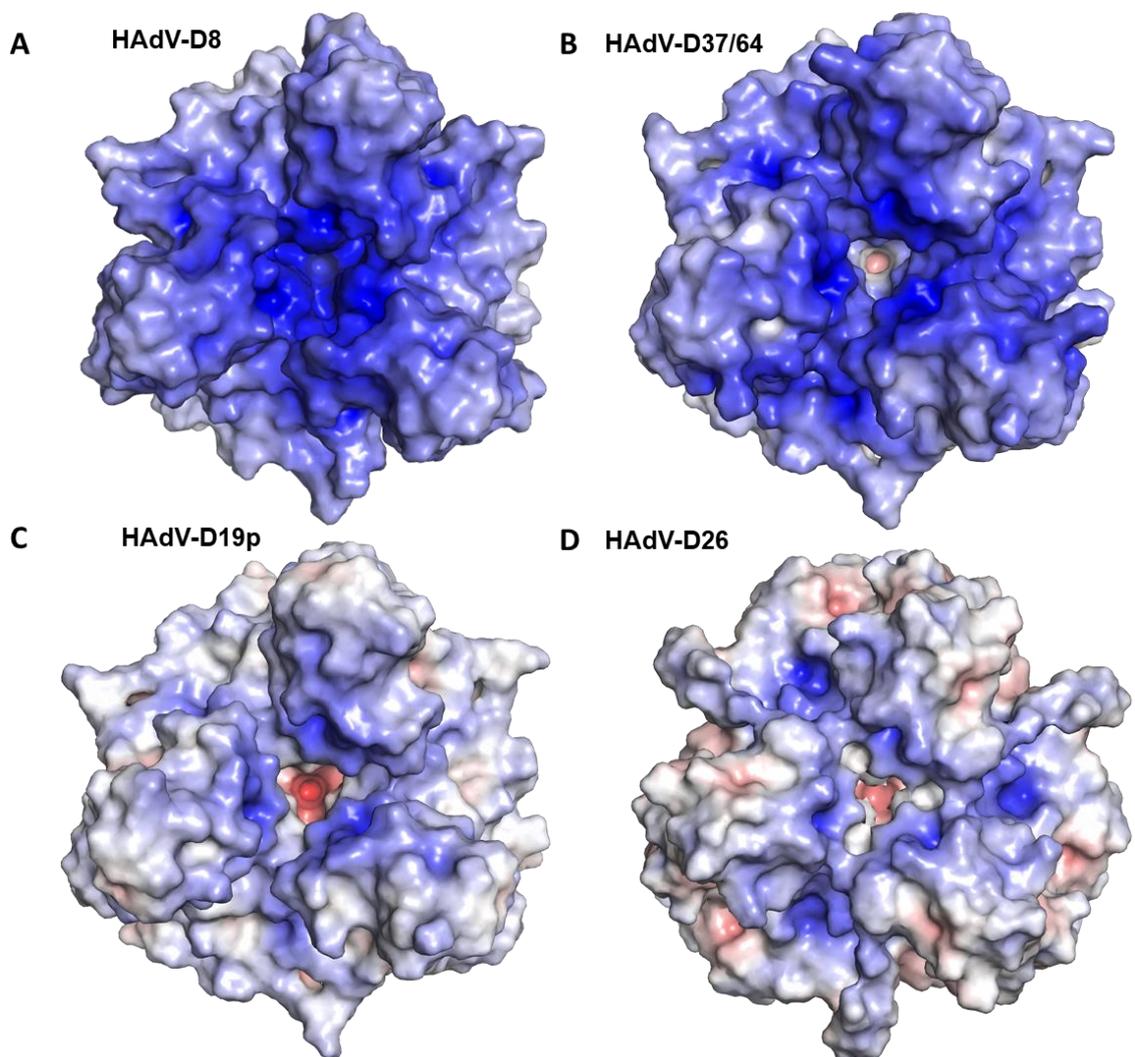


Figure 5-2 Comparison of the electrostatic surface potential of EKC associated adenoviruses. Models of the fiber-knob proteins derived from a homology model of HAdV-D8/53K (**A**), the crystal structure of HAdV-D37/64K (**B**), HAdV-D19pK (**C**), and HAdV-D26K (**D**) are based on crystal structures (1UXA, 1UXB, and 6QU8, respectively) are displayed coloured by electrostatic surface potential. Blue and red region indicate more basic and acidic regions, respectively. Surface potentials were calculated at pH7.35 and the colour ramp is +/- 10mV (blue to red).

D26K is acidic overall with a predicted pI of pH6.49 (Fig5-1B).

We modelled the electrostatic surface potential of the best studied sialic acid utilising adenovirus fiber knob domains, as this has previously been shown to be important to sialic acid binding²⁷, in an attempt to clarify the relationship between the conserved sialic acid residues, the divergent overall sequence, and differing pI's.

As no crystal structure exists for the fiber knob domain of HAdV-D8 we generated a homology model using the closest relative by sequence identity, HAdV-19p, which has an existing crystal structure(Fig5-2A)²⁷. Calculation of the electrostatic surface potentials of HAdV-D8 and HAdV-D37/64 fiber knob domains show an extremely positive overall surface potential (Fig5-2A,B). The fiber knob of HAdV-D19p is less basic (Fig5-2C), in agreement with the predicted pI (Fig-1B). HAdV-D26K shows a similar apical region to HAdV-D19pK, in terms of electrostatic potential, with higher acidity in the lateral regions of the fiber knob (Fig5-2D). Critically, the apical depression of all the modelled fiber knob domains retained basic regions in the apical depression spanning between two monomers, which corresponds to the three sialic acid binding locations previously observed in HAdV-D37 and HAdV-D19p²⁷. This indicates that HAdV-D26K retains the necessary residues and electrostatic potential to form an interaction with sialic acid in a similar manner to known sialic acid interacting adenoviruses.

5.3 HAdV-D26 pseudotyped viruses require sialic acid for efficient cell entry

To test if HAdV-D26 utilises sialic acid as a means of infection we performed infectivity assays treated and untreated with neuraminidase to cleave cell surface exposed sialic acid. We used pseudotyped viruses to ensure the observed effects were due to the fiber knob protein, and not due to interactions with other viral proteins. The tests were performed in SKOV-3 ovarian adenocarcinoma, A549 lung carcinoma, and BT-20 breast carcinoma cells. All three cell lines express CD46 at high levels, A549 cells are high in CAR expression, while BT-20 and SKOV-3 cells only express CAR at a low level.

Our data demonstrate that HAdV-C5/B35K (HAdV-C5 pseudotyped with the fiber knob domain of the CD46 utilising HAdV-B35 virus) infected the three

tested cell lines regardless of neuraminidase treatment, which is unsurprising given the ubiquitous expression of CD46 (Fig5-3A). Similarly, the CAR utilising HAdV-C5 virus retained infectivity following neuraminidase treatment (Fig5-3B). Interestingly HAdV-C5/B35K and HAdV-C5 was more infectious in all tested cell lines, though the effect was more pronounced in HAdV-C5 (Fig5-3A,B). The reasons for this are unclear, but has previously been observed and suggested

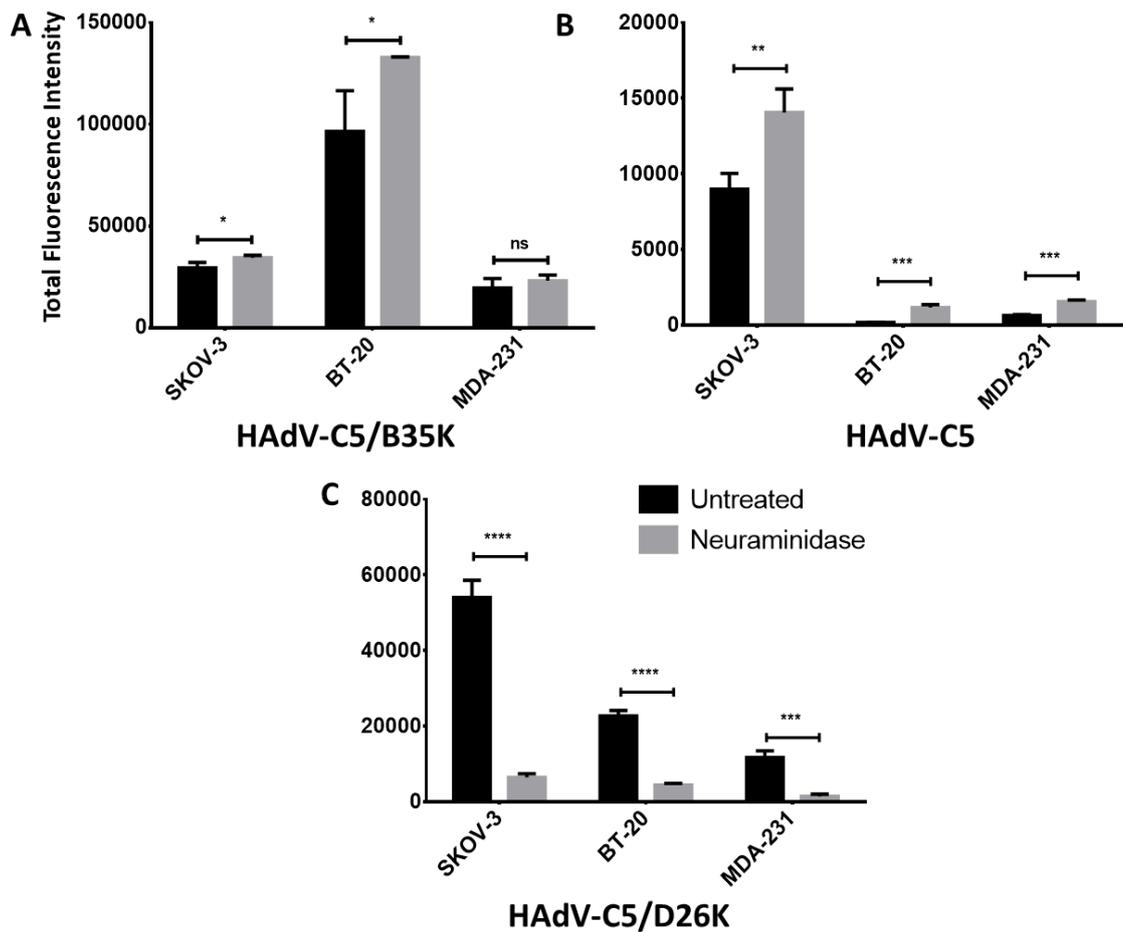


Figure 5-3 Neuraminidase treatment significantly inhibits infection by HAdV-C5/D26K, but not HAdV-C5 or HAdV-C5/B35K, in three cancer cell lines. HAdV-C5/B35K infectivity is unaffected by neuraminidase pre-treatment of SKOV-3, BT-20, or MDA-231 cells (A). HAdV-C5 infectivity increases after neuraminidase treatment of the same cells (B), and HAdV-C5/D26K infectivity is significantly reduced post-neuraminidase treatment (C). $n=3$ biological replicates, error= \pm SD, * $P<0.05$, ** $P<0.001$, *** $P<0.005$, **** $P<0.0001$

to be due to altered electrostatic interactions with the virus capsid retarding cell entry, since the HAdV-C5 capsid is negatively charged^{146,392}.

Infection by the HAdV-C5/D26K virus was significantly inhibited ($P<0.005$) by neuraminidase treatment, resulting in >5-fold inhibition, in all three cell lines tested (Fig5-3C). This indicates that the HAdV-D26 fiber knob forms an

interaction with sialic acid which can be utilised for infection, and that there is a clear preference for sialic acid as the infectious pathway compared to CAR or CD46.

5.4 Crystal structures reveal sialic acid binding in the apex of HAdV-D26K

Table 5-1 Refinement statistics for structures of HAdV-D26K used in this study. One crystal was used for determining each structure. Figures in brackets refer to outer resolution shell, where applicable. MMT (DL-Malic acid, MES monohydrate, Tris: pH 4.0-9.0), SPG (Succinic acid, Phosphate, Glycine). ¹ Coordinate Estimated Standard Uncertainty in (Å), calculated based on maximum likelihood statistics.

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
$\alpha=\beta=\gamma$ (°)	90.0	90.0	90.0
Space group	P 2 ₁ 3	P 2 ₁ 3	P 2 ₁ 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 atoms	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

5.4.1 Sialic acid binding is stable at high and low pH

We obtained crystal structures for HAdV-D26K in complex with sialic acid (Table 1) at pH's between pH4.0 (Fig5-4A-C) and pH8.0 (Fig5-4D-F). Sialic acid

was seen to bind to the same location at both pH4.0 (Fig5-4A) and pH8.0 (Fig5-4D), showing that the interaction is stable throughout the range of pH's. The HAdV-D26K trimer showed very few changes in morphology with only minor variations

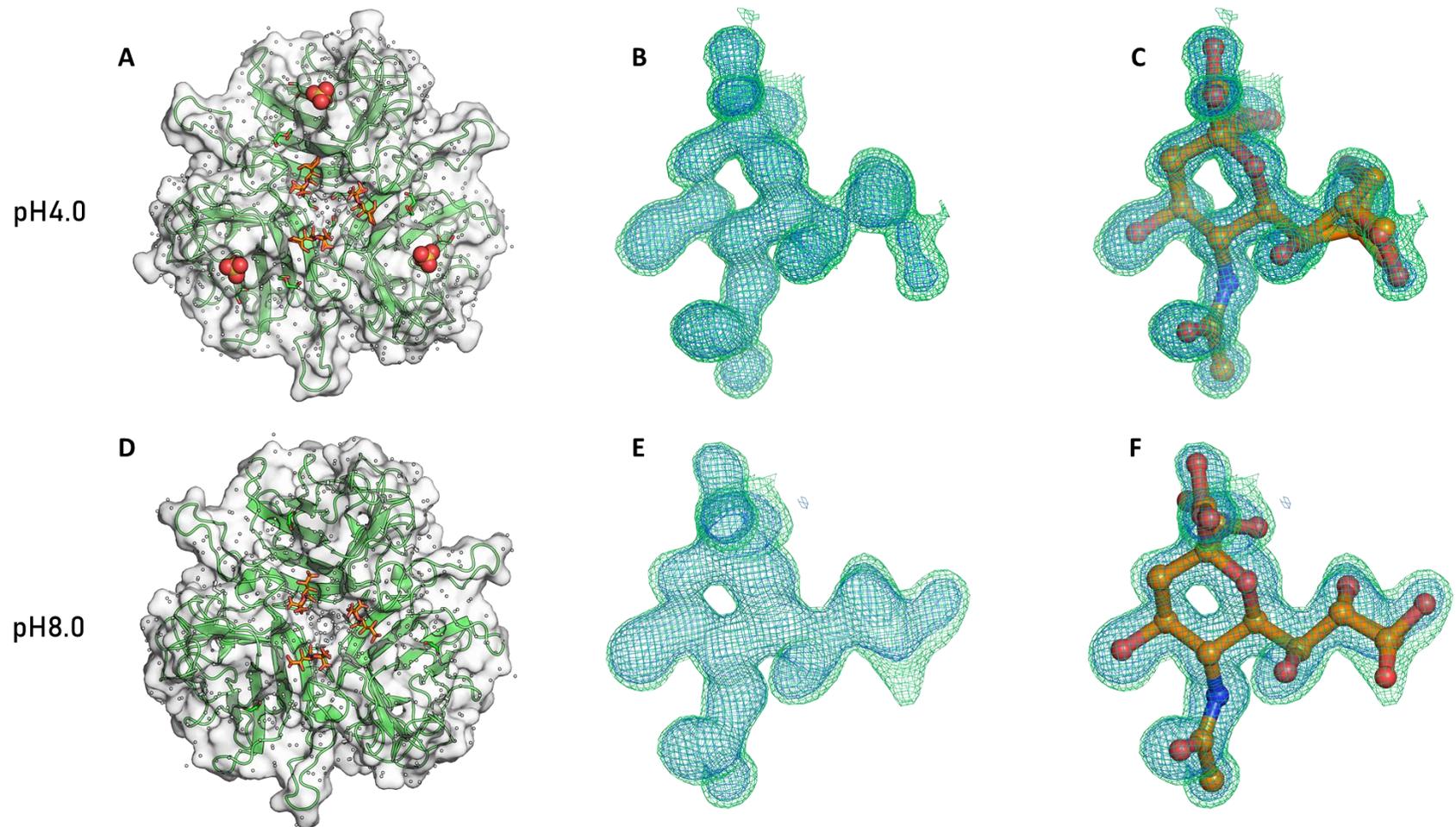


Figure 5-4 Crystal structures show sialic acid interacts with HAdV-D26K in the apical depression at high and low pH. *The crystal structure of HAdV-D26K in complex with sialic acid at pH4.0 (PDB 6QU6) shows three copies of sialic acid interact in the apical depression (A). The omit map shows clear electron density (B) which is best described by a double conformer of sialic acid (C). The same is seen at pH8.0 (PDB 6QU8, D) with electron density (E) showing a different sialic acid double conformer (F).*

inside chain orientation in solvent exposed areas of the loops, showing unusual resilience to changes in acidity.

The omit maps for the sialic acid in the pH4.0 (Fig5-4B) and pH8.0 (Fig5-4E) structures are predominantly similar, but with small differences towards the right of the ring structure. These differences in the map to the right of the ring can be explained best by a double conformer of sialic acid, seen super imposed on the pH4.0 and pH8.0 omit maps, respectively, in figure 4C and 4F. Sialic acid is similar at both pH's with the only significant differences being visible in the glycerol arm. At pH4.0 the arm pivots about the C7-8 bond (Fig5-4C), whereas at pH8.0 the carbons are less mobile, and the C8-9 bond rotates the position of

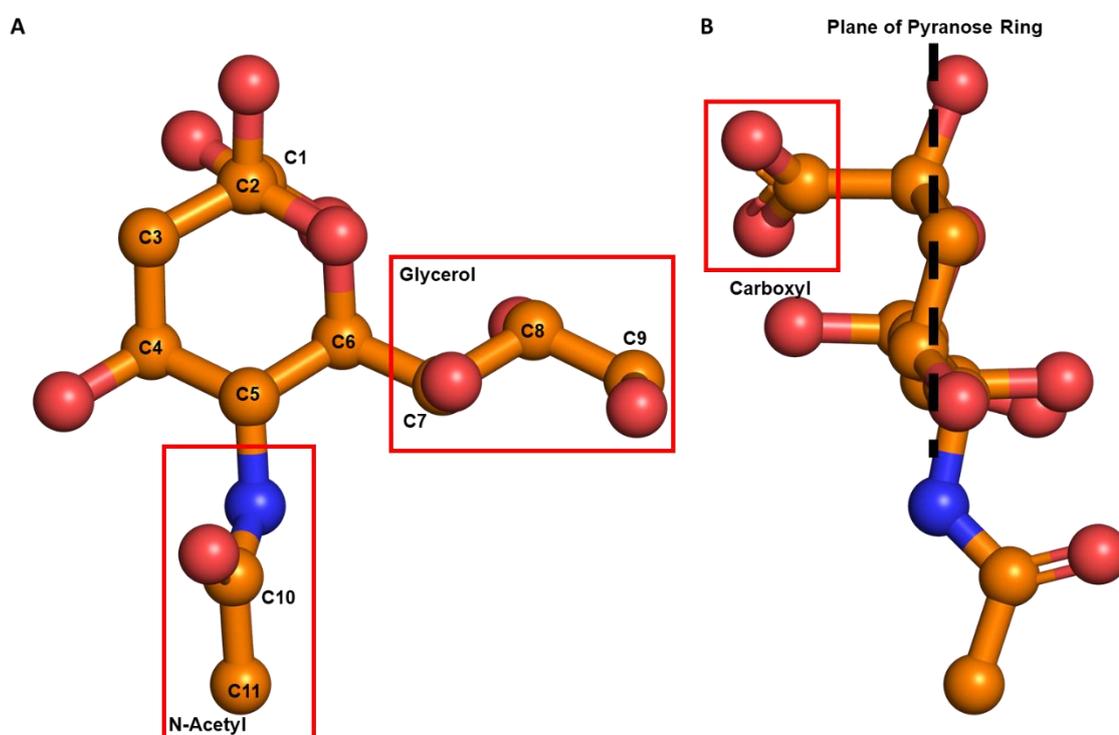


Figure 5-5 The biologically relevant conformation of sialic acid, when in complex with HAdV-D26K. Viewing sialic acid normal to the pyranose ring of sialic acid (A) the glycerol and N-Acetyl functional groups are visible projecting from C6 and C5, respectively. Viewing sialic acid parallel to the pyranose ring the carboxyl group is seen in an axial conformation (B), which is most biologically relevant as it enables access the C2 oxygen group available for α 2 linkages to other sugar monomers.

the OH group (Fig5-4F).

The omit maps show clear density for two orientations of the C1 carboxyl group, both in a planar and an axial conformation (Fig5-4). After refinement the density can be explained with equal occupancy in each conformation, suggesting each are equally valid models of the structure. However, previous work has

demonstrated that in HAdV-D37K^{27,147} and HAdV-D19pK²⁷ the glycan chain extends in the planar direction from C2. This makes the biologically relevant conformation of sialic acid one in which the C2 oxygen is planar (Fig5-5A) and the carboxyl group containing C1 is axial (Fig5-5B).

5.4.2 HAdV-D26K binding to sialic acid utilises multiple types of bond

HAdV-D26K interacts with sialic acid using a combination of polar contacts, water bridges, and hydrophobic interactions at both pH4.0 and pH8.0, showing almost no variation in sialic acid contact residue positions between pH's (Fig5-6). The axial carboxyl group forms polar contacts with a main chain oxygen and a lysine (Fig5-6A,B). While the lysine is seen as a double conformer at pH8.0 (Fig5-6B) it is still in range to form a polar contact. The nitrogen atom of sialic acid also forms a main chain oxygen contact, and the C4 OH group contacts the apical oxygen of a tyrosine side chain. The N-acetyl group forms a polar contact with an asparagine residue, anchoring the sialic acid and orienting the methyl group at C11.

The final polar contacts are seen between a glutamine residue and the closest oxygen of the glycerol arm. The glycerol arm is in a different orientation in each of the two pH's depicted (Fig5-6), though the density evidences that it can form a different pair of conformations at each pH (Fig5-4C,F), suggesting a minimum of 4 possible conformations. However, in all 4 conformations, an oxygen is positioned in the direction of the threonine and in the direction of the glutamine, placing similar polar groups in the same region of space. The oxygen group planar to the ring forms a polar contact with glutamine (Fig5-6A,B), while the oxygen which is oriented towards the threonine forms a water bridge with that amino acid (Fig5-6C,D).

Water bridges are an important mode of interaction for sialic acid binding to HAdV-D26K. At pH4.0, five water bridges are visible (Fig5-6C), and eight are observed at pH8.0 (Fig5-6D). The increased number of water bridges owes to the additional water seen in the pH8.0 structure, contacting between the carboxyl group and a main chain oxygen atom and the side chain nitrogen atom of a lysine residue. Whether or not this water is present in the pH4.0 structure as well is unclear. The pH4.0 structure has higher resolution (1.0Å) compared to

the pH8.0 structure (1.2Å) (Table 1), so we would expect the water bridge to be visible in both. However, it could be that the water is not coordinated enough to form a stable water bridge due to differing protonation states at the lower pH.

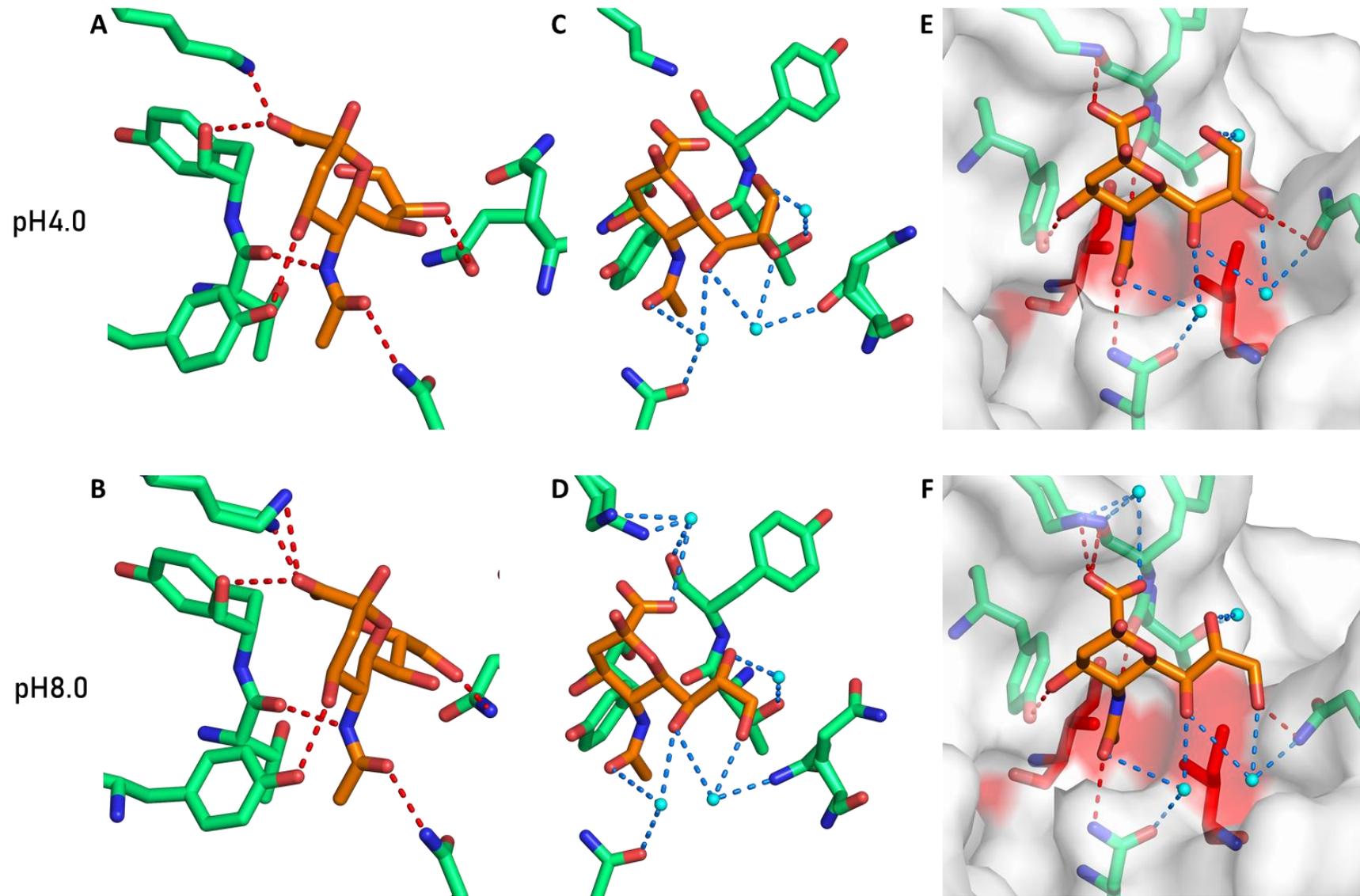


Figure 5-6 HAdV-D26K forms a similar interaction with sialic acid at both pH4.0 (PDB 6QU6) and pH8.0 (PDB 6QU8) through a combination of polar, water-bridge, and hydrophobic interactions. *At pH4.0 sialic acid forms numerous polar contacts to charged side chains in HAdV-D26K (A), similar contacts are 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.*

The final major interaction mode is a hydrophobic interface between a pair of isoleucine residues and the methyl group of the N-acetyl domain of sialic acid. The isoleucine's form a hydrophobic cradle which the methyl group can fit into. As an uncharged mechanism it is unaffected by pH (Fig5-6E, F), however the cyclic face of the tyrosine residue contributes to the hydrophobic interface while also forming a polar contact. This marks an unusual instance of a residue acting as both a polar and hydrophobic agent, simultaneously.

Further inspection of the residues suggests that the sialic acid binding interface in HAdV-D26K may display an induced fit mechanism. Examination of the density by σ -level surrounding the glutamine at position 348, closest to the glycerol arm of sialic acid, indicates that the side chain of Gln348 forms a double conformation. While the main chain is clearly defined up to $\sigma=3.3-37$, at pH4.0 the side chain is shown to be in a conformation contacting sialic acid at $\sigma=0.5-1.2$ (Fig5-7A) and in this conformation at pH8.0 at $\sigma=1.3-1.7$ (Fig5-7B). Weak density is also visible in an alternative confirmation in which the Glu348 side chain is unable to form a polar contact with sialic acid. The Glu348 side chain may have further possible conformations not visible in the density. This suggests that the Glu348 may be labile up until sialic acid binding, then forming a bond with the glycerol arm of sialic acid upon its entry into the binding pocket, stabilising both molecules conformations.

Similarly, Ile324 is seen to have a double conformation for its side chain when not bound to sialic acid (Fig5-7C). One confirmation makes an indentation able to accommodate the methyl group of C11 (Fig5-6E, F), while in the other the long arm of Ile324 rotates to occupy this space. The polar interactions of sialic acid's nitrogen and N-acetyl oxygen hold the methyl in an orientation where it must occupy the hydrophobic cleft formed by Ile324, Ile310, and Tyr312 (Fig5-

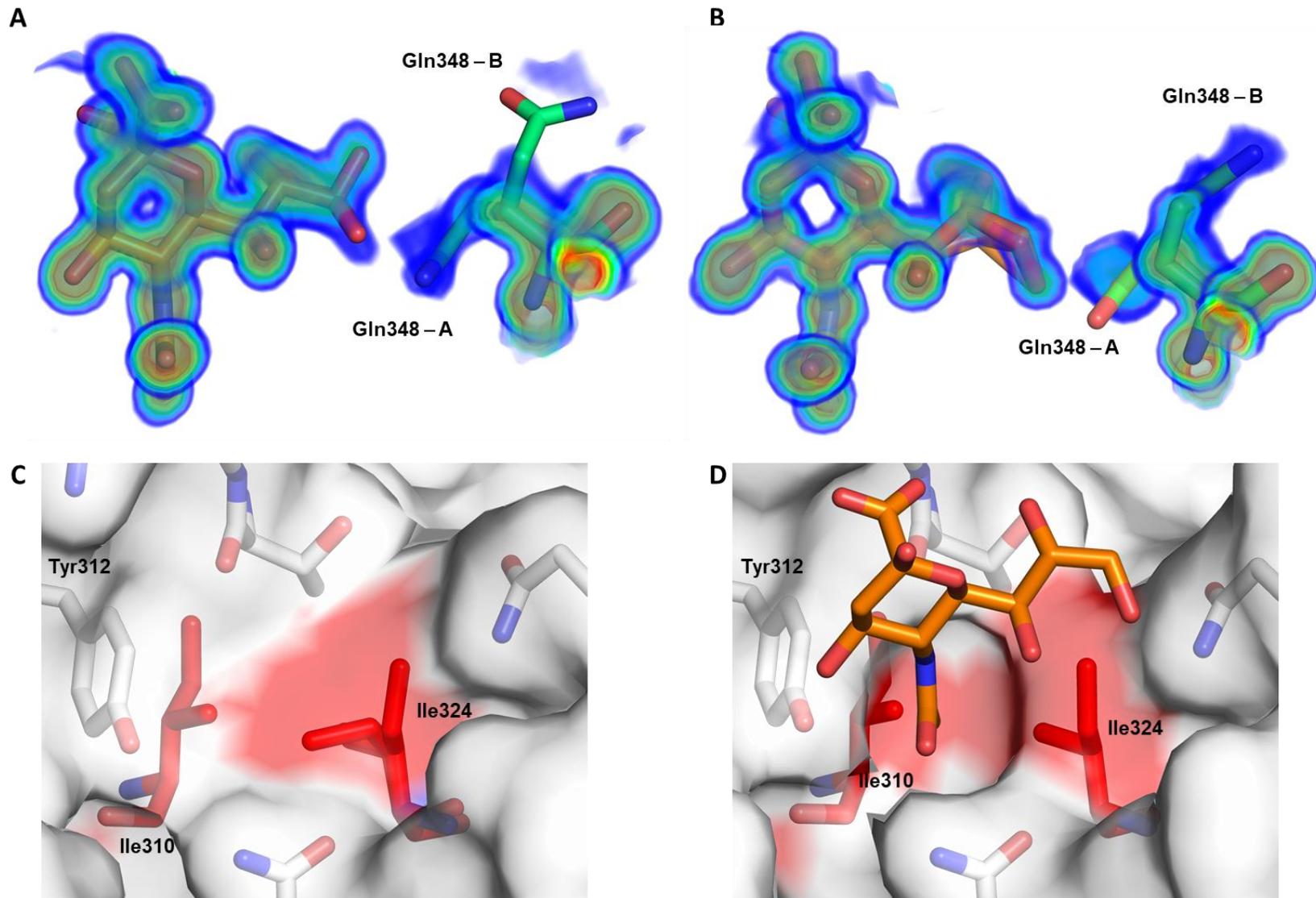


Figure 5-7 HAdV-D26K affects an induced fit mechanism in sialic acid binding. *HAdV-D26K* residue Gln348 can occupy multiple conformations, with a greater preference for conformation A (capable of forming a polar contact with the glycerol arm of sialic acid) at pH8.0 (**A**) than at pH4.0 (**B**). Ile324 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**). Shells describe the σ of the density: Blue: $\sigma=0.5-1.2$, Cyan: $\sigma=1.3-1.7$, Green: $\sigma=1.8-2.2$, Yellow= $2.3-2.7$, Orange: $2.8-3.2$, Red: $\sigma=3.3-3.7$.

6A, B; Fig5-7D). It is unclear whether the Ile324 must already be in a permissive conformation for sialic acid interaction, or if it can be “pushed” into this side chain orientation by sialic acids other interactions. Regardless of the order in which this occurs, it is an example of the conformation of sialic acid and HAdV-D26K morphing into a shape permissive for an interaction, and thus is an example of induced fit.

5.5 HAdV-D26 binds to sialic acid in a manner similar to known sialic acid utilising species D adenoviruses

The interaction of HAdV-D26K with sialic acid is similar to that of HAdV-D37K and HAdV-D19pK. Figure 8 shows a comparison of HAdV-D26K and HAdV-D37K in complex with sialic acid. Both demonstrate a similar overall mechanism of binding sialic acid, with the N-acetyl domain of sialic acid anchored to a polar residue, the methyl group contacting a hydrophobic patch, and the nitrogen, carboxyl, and C4 OH group restrained by direct polar interactions. However, HAdV-D26K has a more numerous network of bonds which restrain the sialic acid residue more tightly. How this affects the binding affinity is unclear, though it seems likely to increase the affinity of the interaction, though the steric interactions with the isoleucine arms could diminish the on rate.

The carboxyl group of both HAdV-D26K and HAdV-D37K is restrained by a direct polar contact to Lys349/Lys345 respectively, while the HAdV-D26K carboxyl group is further restrained by a water bridge and another direct polar contact with the main chain oxygen of Tyr320 (Fig-8A). In HAdV-D37K the oxygen of the N-acetyl group is bound by a pair of water bridges to the oxygen of Thr310 (Fig-8B). In HAdV-D26K Thr310 is substituted for an asparagine residue (Asn312, Fig-8A). The Asn312 side chain is longer, and further differs from threonine by having both a positively charged oxygen, and negatively charged nitrogen group. This enables the formation of both a direct polar interaction, and a water bridge. The same water is close enough to the C7 OH group to form another water bridge (Fig-8A).

The glycerol arm of HAdV-D37K is restrained by a single water bridge between the C7 OH and Ser344 (Fig-8B). The serine side chain is too short to form a direct polar interaction. The equivalently positioned residue to Ser344 in HAdV-D26K is Gln348. Much like the Thr310-Asn312 difference, Gln348 is also

significantly longer than its equivalent residue in HAdV-D37K and has both

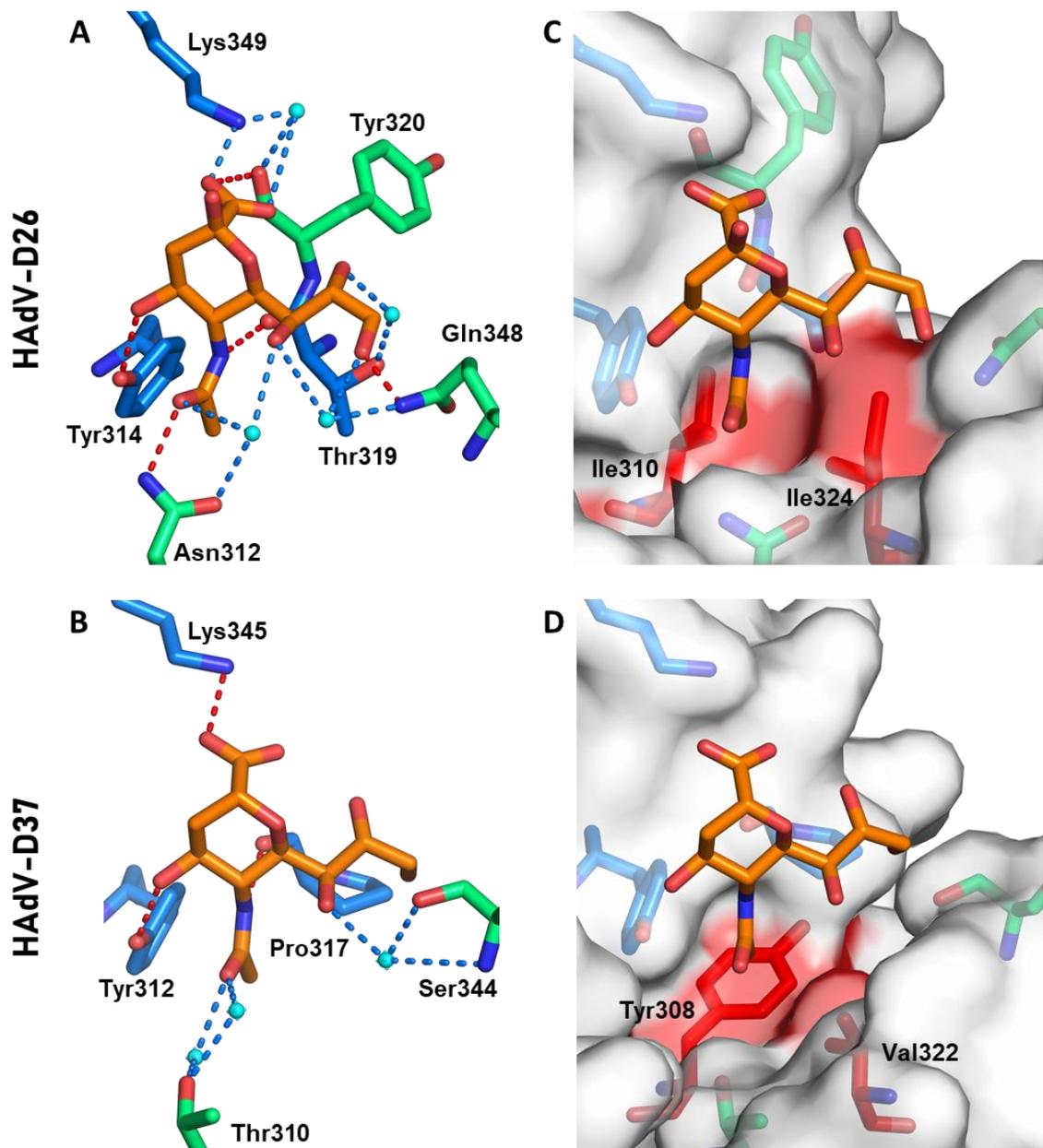


Figure 5-8 HAdV-D26K forms a complex interaction network of hydrophobic and electrostatic interactions with sialic acid. Sialic acid (orange) is seen to bind HAdV-D26 (**A**) and HAdV-D37 (**B**) through a network of polar contacts (red dashes) and hydrogen bonds (blue dashes). The interaction is stabilised by hydrophobic interactions (red regions on white surface) with the N-Acetyl CH3 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.

positive and negative charges. Gln348 forms a direct polar interaction to the C9

OH group of sialic acid, and further stabilises the glycerol arm by water bridge contacts to both C7 and C9 OH groups (Fig5-8A). Finally, Thr319 forms a water bridge to the C8 OH group. By these mechanisms HAdV-D26K appears to restrain sialic acid more tightly than HAdV-D37K, forming contacts at 8 points, rather than 5, and many more bonds overall.

The “cradle” conformation formed by Ile310 and Ile324 in HAdV-D26K is another example of sialic acid being more restrained than in HAdV-D37K (Fig-8C). In HAdV-D37K the equivalent interface is composed of Tyr308 and Val322, forming a flat hydrophobic patch (Fig5-8D). This flat region provides empty space which the sialic acid could move into, in contrast to the cradle formed in HAdV-D26K which is likely to be further controlled by the pressure from the labile Ile324 side chain (Fig5-7C, D).

5.6 Species D adenoviruses conserve sialic acid binding features

Recent research has confirmed that the majority of the EKC causing adenoviruses utilise sialic acid as a cell entry receptor, which follows, given their shared pathology and high sequence similarity, including key sialic acid binding residues⁵⁹. However, the above finding that HAdV-D26 utilises sialic acid as a receptor suggests that this tropism may not be restricted to the EKC causing adenoviruses and may exist more broadly.

Amino acid sequence alignment of all the known species D adenoviruses shows conservation of many residues known to be involved in sialic acid interaction (Fig5-9). HAdV-D26K has isoleucine residues at positions 310 and 324 which correspond to tyrosine in previously known sialic acid binding viruses HAdV-D37/19p/8/64. In other species D fiber knobs position 310 is typically valine, while 324 is poorly conserved. In HAdV-D26K residue 312 is asparagine, and threonine in HAdV-D37K and HAdV-D19p. However, the asparagine is the common, being conserved at that position in the entire species D except for HAdV-D10K/37K/19pK Lys349 is also conserved across the entire species D, similarly Tyr314.

We have generated crystal structures of HAdV-D10K (Fig5-10), D15K (Fig5-11), D29K (Fig5-12), and D30K (Fig3-14). HAdV-D10K packed as a pair of biological

trimers in P6₃ at a resolution of 3.39Å (Fig5-10, Table5-2). This low resolution would be difficult to solve in the absence of a close pre-existing model, but molecular replacement with the very closely related structure of HAdV-D37K

Table 5-2 Refinement statistics for deposited structures of previously undescribed species D adenoviruses. One crystal was used for determining each structure. Figures in brackets refer to outer resolution shell, where applicable. SPG (Succinic acid, Phosphate, Glycine). ¹ Coordinate Estimated Standard Uncertainty in (Å), calculated based on maximum likelihood statistics.

enabled successful phasing. We were able to define the main chain of the

PDB Entry	HAdV-D10K – 6QPM	HAdV-D15K – 6STW	HAdV-D29K – 6STV
Data Collection			
Diamond Beamline	IO4-1	I03	I03
Date	2018-03-12	18/04/2019	18/04/2019
Wavelength	0.91587	0.95372	0.95372
Crystallization Conditions	0.2 M Na NO ₃ , 0.1 M Bis-Tris Propane, 20.0 % PEG 3350	0.1M SPG, 25% w/v PEG 1500	0.1M SPG, 25% w/v PEG 1500
pH	7.5	8.0	6.0
<i>a, b, c</i> (Å)	183.57, 183.57, 94.88	59.58, 89.18, 106.20	104.29, 104.29, 104.29
α, β, γ (°)	90.0, 90.0, 120.0	90, 90, 90	90, 90, 90
Space group	P 6 ₃	P 2 ₁ 2 ₁ 2 ₁	P 2 3
Resolution (Å)	3.395 – 79.49	1.37 – 39.64	1.60 – 73.85
Outer shell	3.39 – 3.58	1.37 – 1.41	1.60 – 1.63
<i>R</i> -merge (%)	10.0 (245.5)	0.036 (1.604)	0.105 (3.525)
<i>R</i> -meas (%)	10.4 (256.9)	0.041 (1.862)	0.109 (3.690)
CC1/2	0.999 (0.426)	1.000 (0.478)	1.000 (0.444)
<i>I</i> / σ (<i>I</i>)	15.5 (1.0)	18.4 (1.1)	18.1 (1.0)
Completeness (%)	99.6 (97.1)	100.0 (100.0)	100.0 (100.0)
Multiplicity	11.3 (11.3)	7.4 (7.6)	22.6 (22.5)
Total Measurements	285,257 (40,066)	881,087 (66,057)	1,131,680 (54,939)
Unique Reflections	25,265 (3,551)	119,296 (8,721)	50,042 (2,447)
Wilson B-factor(Å ²)	118.1	19.9	24.2
Total number of refined atoms	8,631	5,429	3,247
R-work reflections	23,861	113,301	47,543
R-free reflections	1,235	5,905	2,470
R-work/R-free (%)	20.3 / 23.4	17.3 / 19.9	17.6 / 20.1
Bond lengths (Å)	0.008	0.012	0.012
Bond Angles (°)	1.695	1.722	1.699
¹ Coordinate error	0.451	0.049	0.075
Mean B value (Å ²)	170.2	27.8	30.0
Favoured/allowed/Outliers %	974 / 96 / 6 90.5 / 8.9 / 0.6	451 / 27 / 4 93.57 / 5.60 / 0.83	317 / 13 / 0 96.06 / 3.94 / 0

structure with confidence; however, the specific side chain positions are only

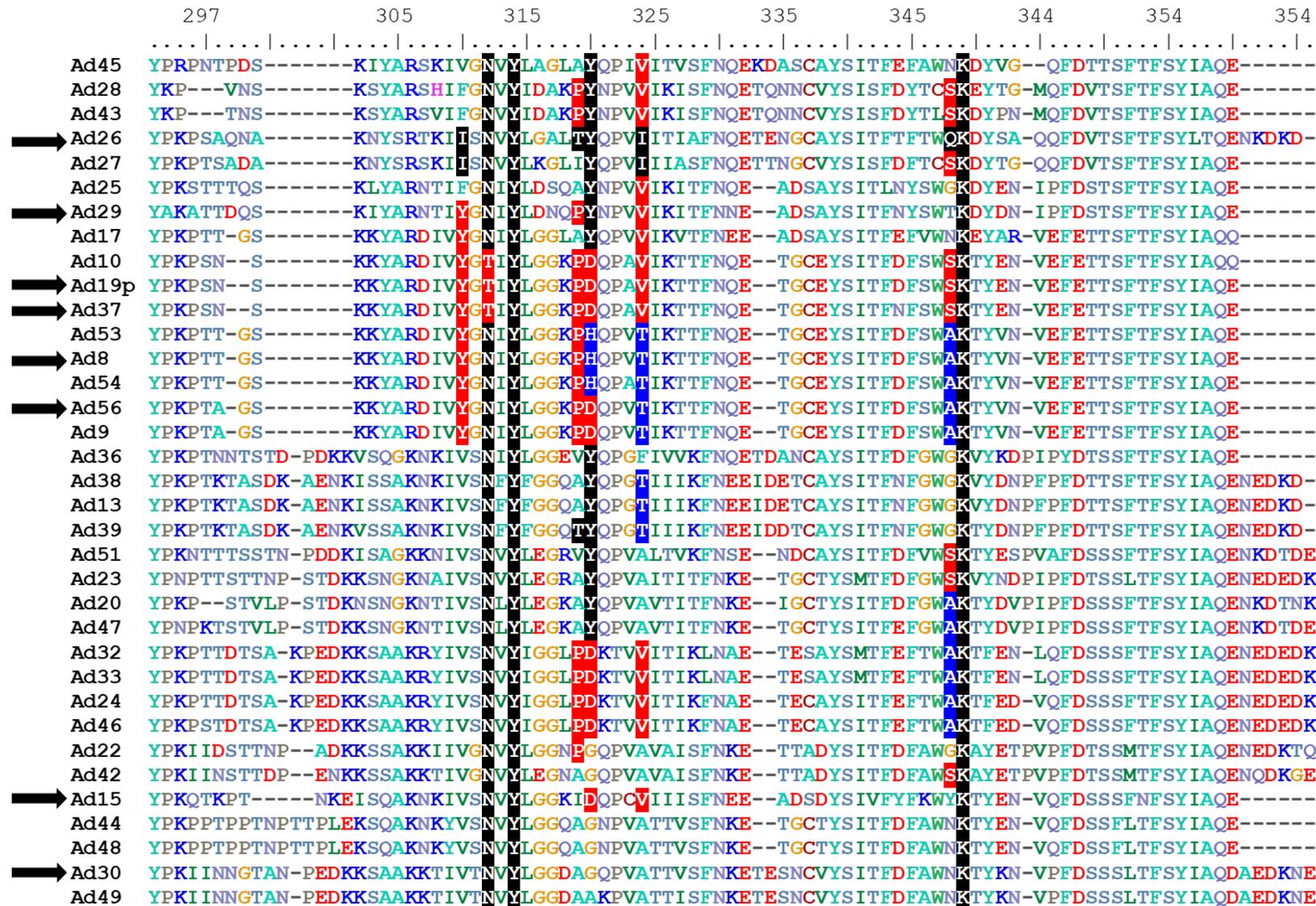


Figure 5-9 Species D adenoviruses conserve known sialic acid binding residues. Sequence alignment of species D adenovirus fiber-knob proteins. Known HAdV-D26 and HAdV677 D37/64/19p residues forming contacts with sialic acid are highlighted in black and red respectively. Homologous residues are coloured 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 for readability, all are human species D adenoviruses. Numbering is based on the sequence of HAdV-D26K. Arrows indicate fiber knob domains with known, or potential sialic acid binding ability.

defined with confidence in higher quality map regions (Fig5-10B). The statistics were commensurate with other structures of this resolution (Fig5-10C).

HAdV-D15K packed as a biological trimer per asymmetric unit (Fig5-11A). The high resolution (1.37Å) provides excellent density enabling us to define side chain positions confidently (Fig5-11B, Table5-2). The statistics were good in comparison to other structures at similar resolution (Fig5-11C).

HAdV-D29K formed a crystal structure in space group P 2 3, containing two copies of the monomer per asymmetric unit, each forming a separate trimer when the symmetry copies are viewed (Fig5-12A). Density was seen to a resolution of 1.60Å and enabled clear definition of most residues (Fig5-12B), and generated statistics were acceptable quality. However, no density was visible for 6 residues contained within a loop at the base of the HAdV-D29K monomer, seen in Fig5-12A, as such these residues were not modelled.

Using the species D fiber knob crystal structures we investigated the potential to interact with sialic acid. Sequence alignment (Fig5-9) identified the amino acids at equivalent positions to known sialic acid contact residues in HAdV-D26K and HAdV-D37K. These residues were compared using the non-ligand bound structures to avoid visualising the effects of induced fit conformations for HAdV-D26K and HAdV-D37K in complex with sialic acid (Fig5-13).

This comparison revealed, as would be expected from the sequence alignment, that HAdV-D10K has an identical sialic acid binding pocket to HAdV-D37K. HAdV-D29 shares most of the residues in this pocket with HAdV-D26K, but with the two hydrophobic residues (Tyrosine and Valine) being the same as those in HAdV-D37K. These similarities suggest that HAdV-D10K and HAdV-D29K are likely to interact with sialic acid.

HAdV-D30K, 48K, and 49K have nearly identical pockets to each other (Fig5-13), the only difference being HAdV-D49K has an Alanine rather than a Guanine at position 320 (Fig5-9) as occurs in HAdV-D30K and 48K. However amino acid changes at this position are unlikely to affect sialic acid binding as the side chain faces out of the binding pocket and the contact is formed with the main chain oxygen. The high degree of similarity to known sialic acid binding pockets suggest that these three fiber knobs may be able to interact with sialic acid.

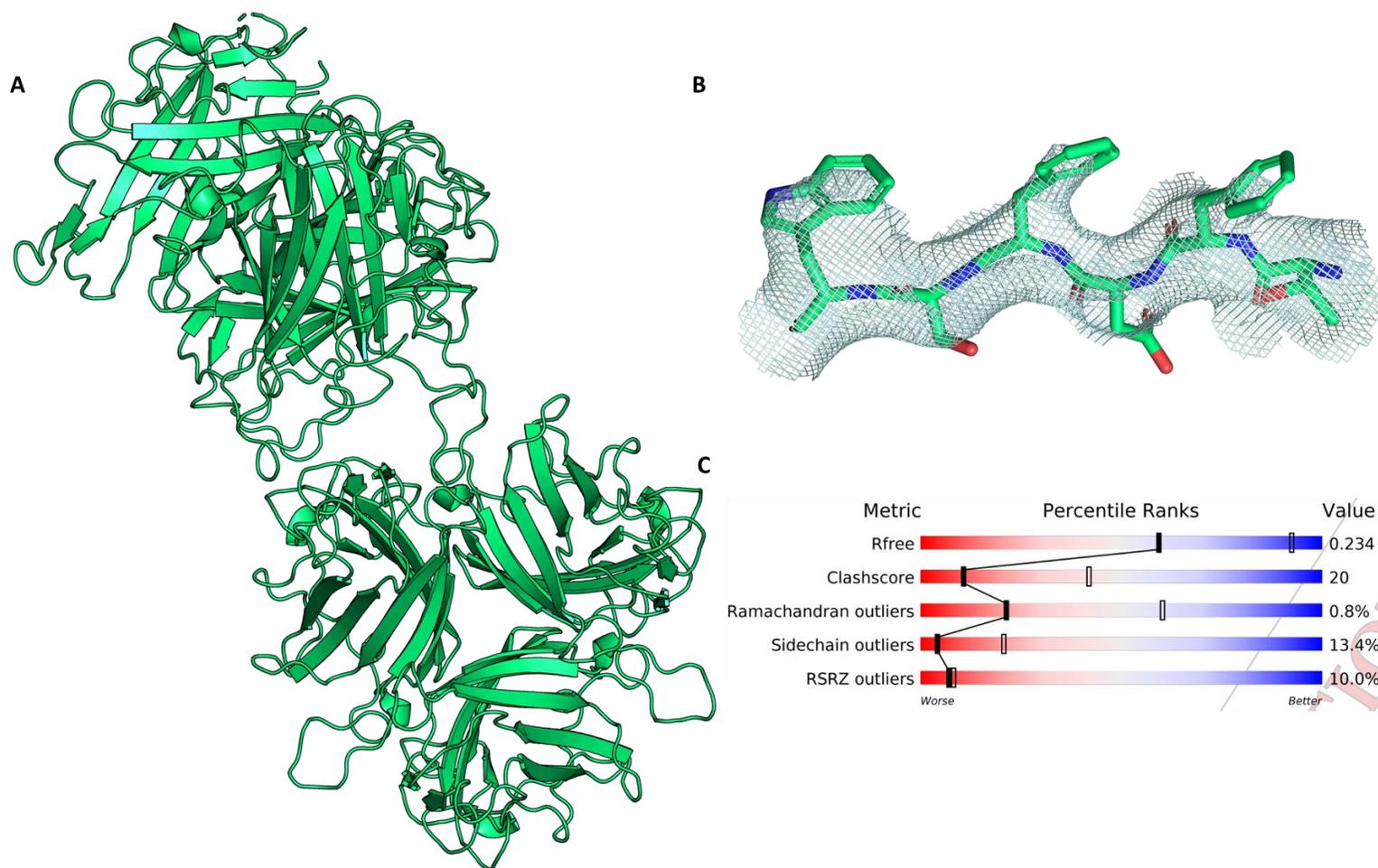


Figure 5-10 Overview of the solved crystal structure of HAdV-D10K, deposited in the wwPDB as 6QPM. The asymmetric unit consisted of 6 fiber-knob monomers, creating a pair of trimers, representing two biological units (A), the quality of the map (blue mesh) through the main β -strands was sufficient to suggest likely side-chain positions (B). The statistics showed scores commensurate with other wwPDB entries at this resolution (C). Map is displayed at $\sigma=1.0$, and carved at 2.0\AA . Protein is shown as a cartoon representation, with missing portions shown as dashed loops.

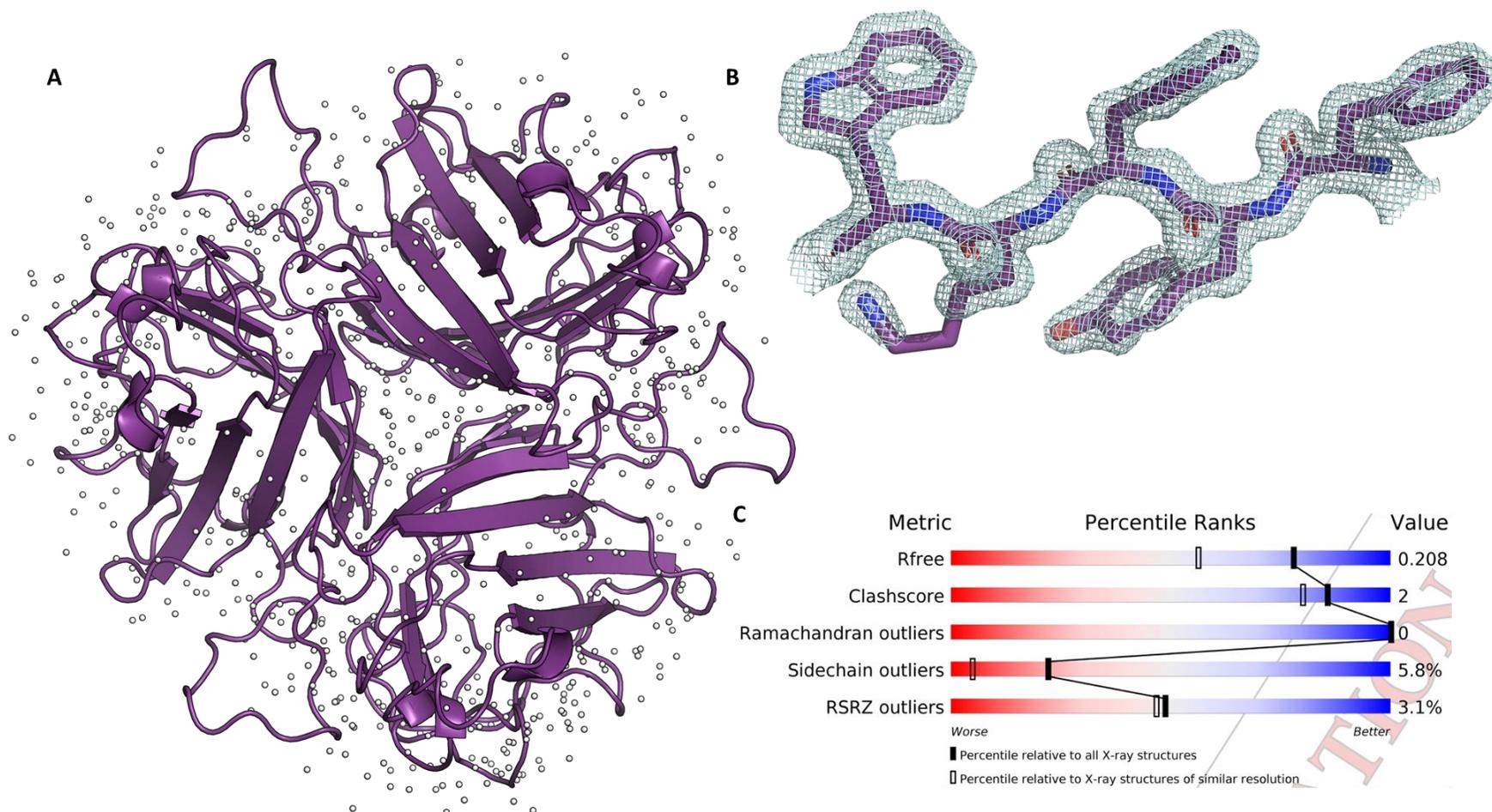


Figure 5-11 Overview of the solved crystal structure of HAdV-D15K, deposited in the wwPDB as 6STW. The asymmetric unit consisted of 3 fiber-knob monomers, creating a trimer, representing the biological unit (A), the quality of the map (blue mesh) through the main β -strands was sufficient to clearly define the side chain positions and their orientation (B). The statistics showed scores commensurate with other wwPDB entries at this resolution (C). Map is displayed at $\sigma=1.5$ and carved at 1.5Å. Protein is shown as a cartoon representation, with water molecules shown as white non-bonded spheres.

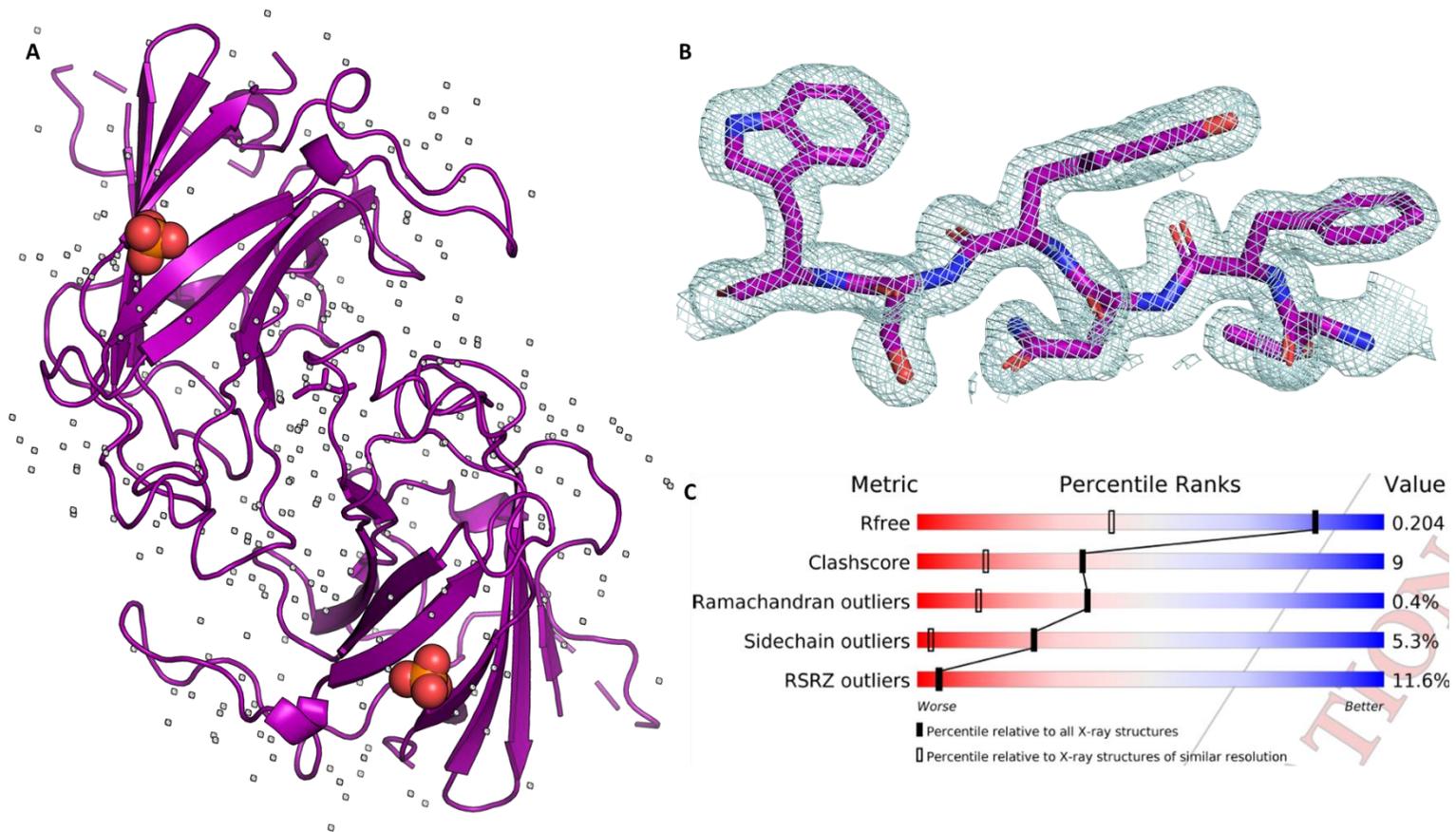


Figure 5-12 Overview of the solved crystal structure of HAdV-D29K, deposited in the wwPDB as 6STV. The asymmetric unit consisted of 2 opposing fiber-knob monomers (**A**), the quality of the map (blue mesh) through the main β -strands was sufficient to clearly define the side chain positions and their orientation (**B**). The statistics showed scores commensurate with other wwPDB entries at this resolution (**C**). Map is displayed at $\sigma=1.5$ and carved at 2.0\AA . Protein is shown as a cartoon representation, with water molecules shown as white non-bonded spheres, and phosphate atoms as Van der Waal spheres. Missing portions of structure are shown as dashed lines.

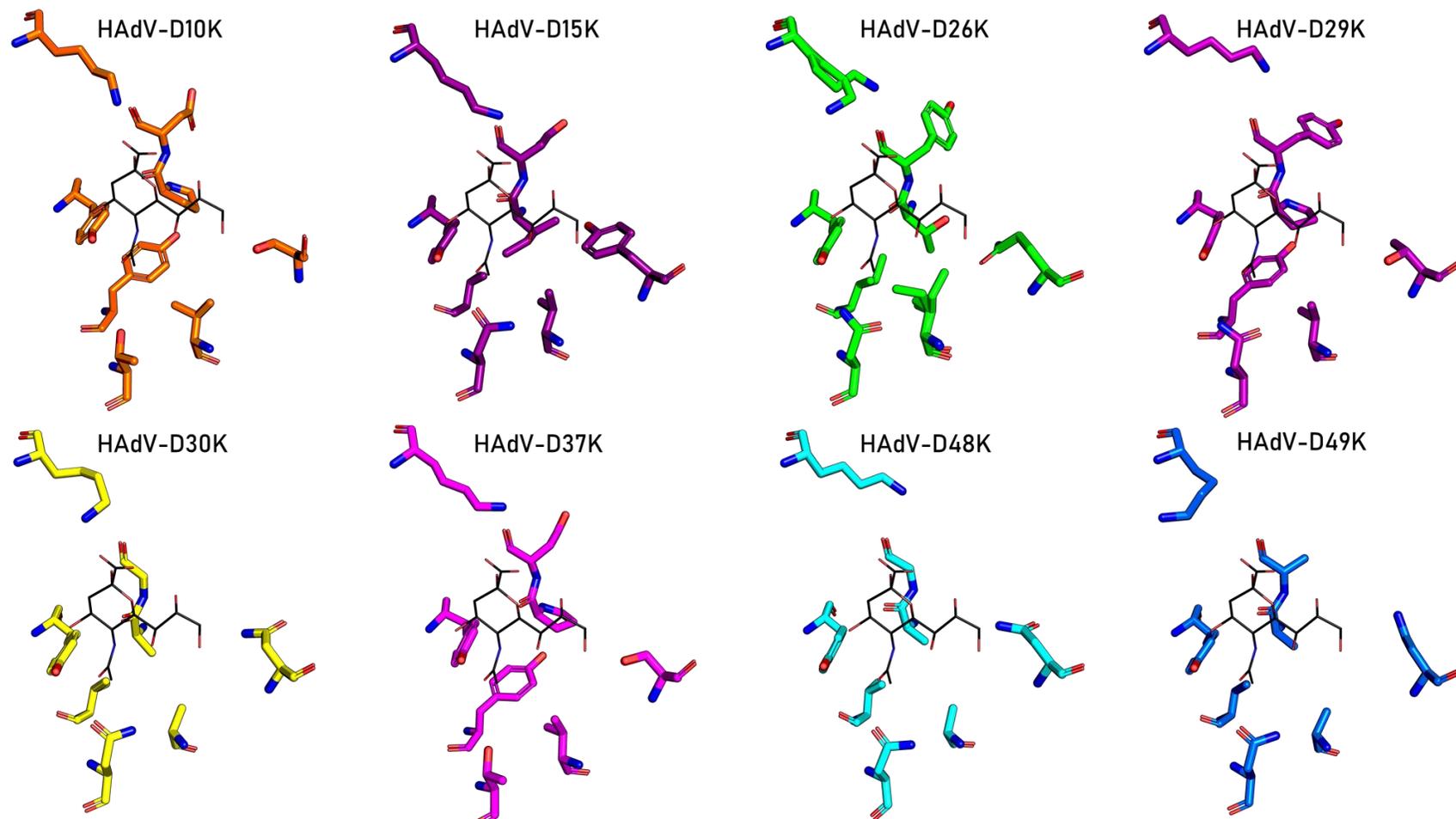


Figure 5-13 Sialic acid binding residues of HAdV-D26K and the equivalent residues, as defined by amino acid sequence alignment, in other species D adenovirus crystal structures. *The sialic acid binding residues of HAdV-D26K and the residues at equivalent positions as defined by ClustalΩ amino acid sequence alignment of the fiber-knob domains are shown. Sialic acid (black lines) is placed by homology alignment to the unliganded structures, and is only present as a guide to assist with orientation, not as a model of interaction.*

HAdV-D15K shares the Asp312 and Tyr314 residues found in HAdV-D26K, as well as the Lys349 found through the species D fiber knobs. The hydrophobic region is also conserved, consisting of a pair of valine residues. HAdV-D15K is dissimilar to other fiber knobs compared in figure 5-13 as the residues at position 319 is a large hydrophobic (isoleucine) which faces into the pocket in which the N-Acetyl methyl group of sialic acid interacts with the fiber knob in HAdV-D26K and 37K. HAdV-D15K Tyr348 is also different in that it is a much larger side chain compared to the serine and glutamine found in HAdV-D37K and 26K, respectively. While the overall character of the pocket is similar to that of the known sialic acid binding fiber knobs, the blocking of the hydrophobic pocket for the N-acetyl methyl group and occupancy of the area required for the glycerol arm of sialic acid suggest HAdV-D15K is unlikely to bind to sialic acid.

Using existing crystals of HAdV-D15K, D29K, D30K we performed soaking experiments to investigate their ability to bind sugars. Since sialic acid is not the only potential sugar monomer on glycans which the fiber knob proteins could, potentially, bind to, we tested binding to a mixture of sialic acid (Neu5Ac), Mannose, Galactose, Fucose, and N-acetyl Glucosamine.

In HAdV-D29K we observe evidence of ligand binding. Density was visible for an N-Acetyl domain, and weaker density extended from a C6 suggesting an “arm” in both likely sialic acid binding sites observed in the soaked crystal structure (Fig5-14A, B). Since sialic acid is the only ligand soaked into the structure with both a large extended functional group (“arm”) and an N-acetyl domain is sialic acid, it is the only explanation for the density. Sialic acid is seen to bind HAdV-D29K in much the same manner as it does to HAdV-D26K. The best fit for the electron density places one arm lateral to the plane of the pyranose ring (Fig5-14A) and the other copy under torsion (Fig5-14B). However, the glycerol arm has weak density suggesting poor localisation regardless of the refined orientation.

Solving the structure of the soaked HAdV-D30K crystal showed density for a small molecule ligand in the region expected for sialic acid binding (Fig5-14C). However, the density in this region was poorly defined, this may be due to insufficient resolution of the data set obtained (2.56Å) or partial occupancy at the site due to insufficient soaking duration. While a ligand can bind in this

location, attempting to fit the density with Mannose, Galactose, or Sialic acid, all gave a plausible

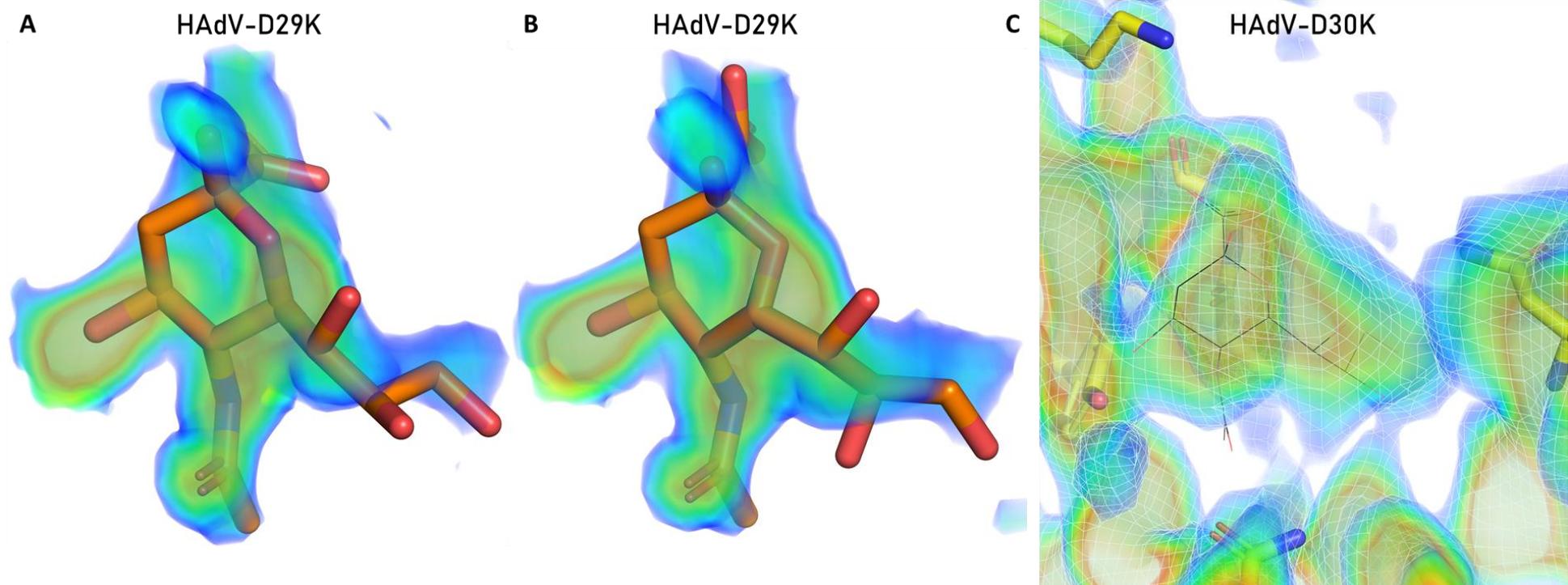


Figure 5-14 Data collected from small molecule soaked HAAdV-D29K crystals shows density corresponding to sialic acid binding while similar data from HAAdV-D30K shows non-descript small molecule binding. Two copies of HAAdV-D29K monomer are present in the asymmetric unit of the crystal structure of HAAdV-D29K. Density corresponding to the shape of sialic acid was visible in both, with one copy placing sialic acid under less torsional stress (**A**) than the other (**B**) and both showing weak density for the glycerol arm. HAAdV-D30K shows a large region of undefined density in the region of the soaked HAAdV-D30K crystal structure but does not have sufficient resolution to define features (**C**).

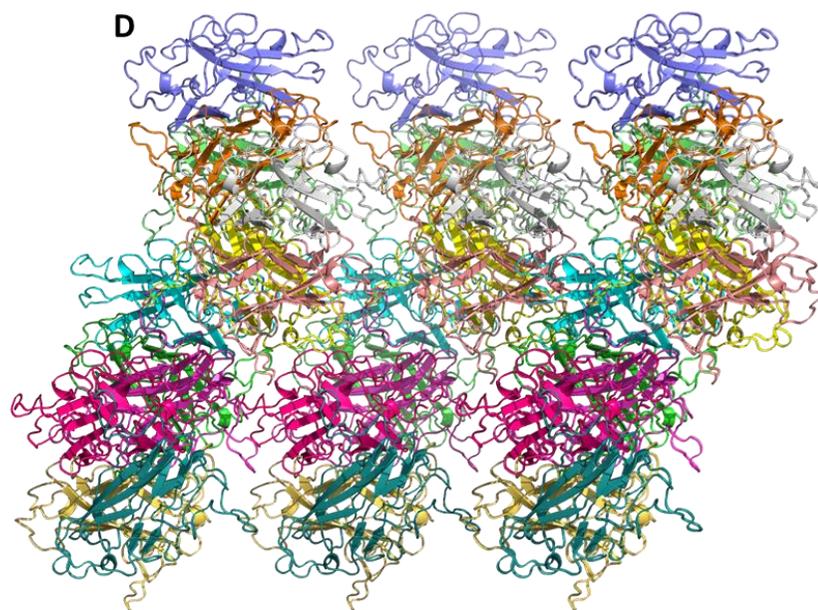
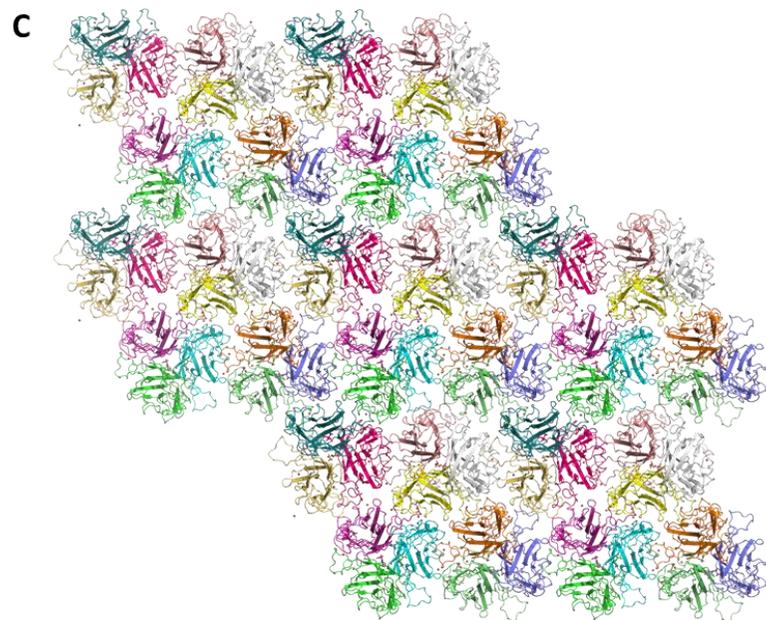
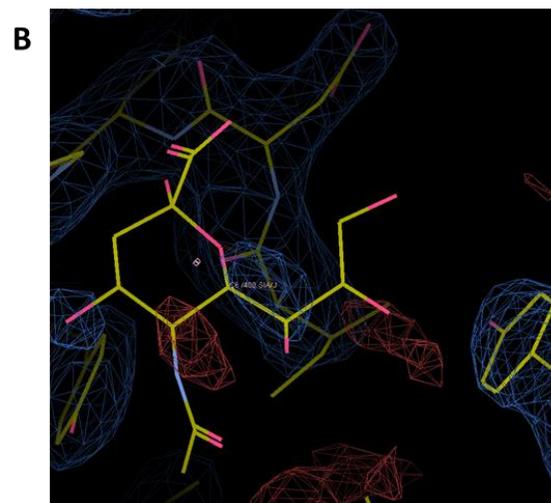
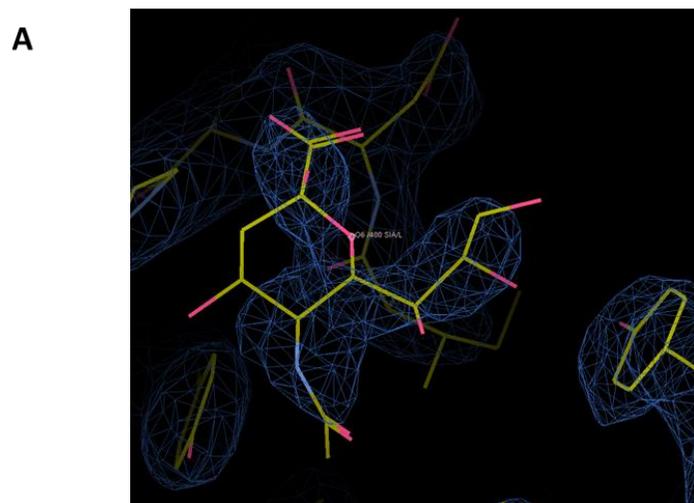


Figure 5-15 Soaked structure of HAdV-D15K does not show clear sialic acid binding density, and has a very tight crystal packing arrangement.

The data set obtained from a small molecule soaked HAdV-D15K crystals included 12 copies of the HAdV-D15K monomer for 4 biological trimer and 12 possible sialic acid binding sites in the apical domain. Some potential sites showed weak ($\sigma < 1.0$) density in regions of the possible sialic acid binding pocket in some copies (A) and either no density or negative density in other copies (B). The crystal packing arrangement showed a very close packing arrangement with no clear diffusion channels running parallel to the 3-fold axis (C) and large loops blocking access to the apical domains of the biological trimers (D).

explanation for the density at all three of the potential bind sites, leaving the identity of the ligand inconclusive.

Soaking experiments in HAdV-D15K showed no evidence of sugar binding, as suggested by earlier modelling of this binding pocket (Fig5-13). The structure of HAdV-D15K which was solved had four copies of the HAdV-D15K trimer, and thus twelve possible binding sites for sialic acid. Adding sialic acid by homology and re-refining the data set showed some density around one potential site of sialic acid binding (Fig5-15A). However, most sites showed no density (Fig5-15B). The packing arrangement of the soaked HAdV-D15K structure exhibited a particularly close packing arrangement with small channels in the direction of the trimers three-fold axis (Fig5-15C) and loops extending into the trimer's apex, across the plane normal to the three-fold axis (Fig5-15D). This packing arrangement could inhibit diffusion through the crystal.

We further assessed the biological significance of sialic acid presence on the cell surface for viral infection by performing luciferase infectivity assays and treating with neuraminidase (Fig5-16). We tested viral pseudotypes HAdV-C5/D10K, HAdV-C5/D15K, HAdV-C5/D24K, HAdV-C5/D29K, HAdV-C5/D53K, HAdV-C5/D26K. As in the flow cytometry assay, HAdV-C5/D26K was significantly inhibited by neuraminidase treatment in SKOV-3 cells and is now shown to depend upon sialic acid to infect A549 lung carcinoma cells as well (Fig-16A). HAdV-C5/D10K was significantly less capable of infecting A549 cells after neuraminidase treatment, and while a reduction in infectivity was seen in SKOV-3 cells as well it was not significant (Fig5-16B). This pattern was also seen in HAdV-C5/D15K, and may be explained by rotation of the isoleucine seen to inhibit the hydrophobic interface being possible in a non-crystallised structure (Fig5-16C). A reduction in infectivity was seen for HAdV-D24K in A549 cells, but neither this nor the effect in SKOV-3 cells was significant. HAdV-C5/D29K followed the pattern observed in HAdV-C5/D10K and D15K, while HAdV-C5/D53K showed the opposite with a significant reduction in infectivity observed in SKOV-3 cells but a non-significant reduction in A549.

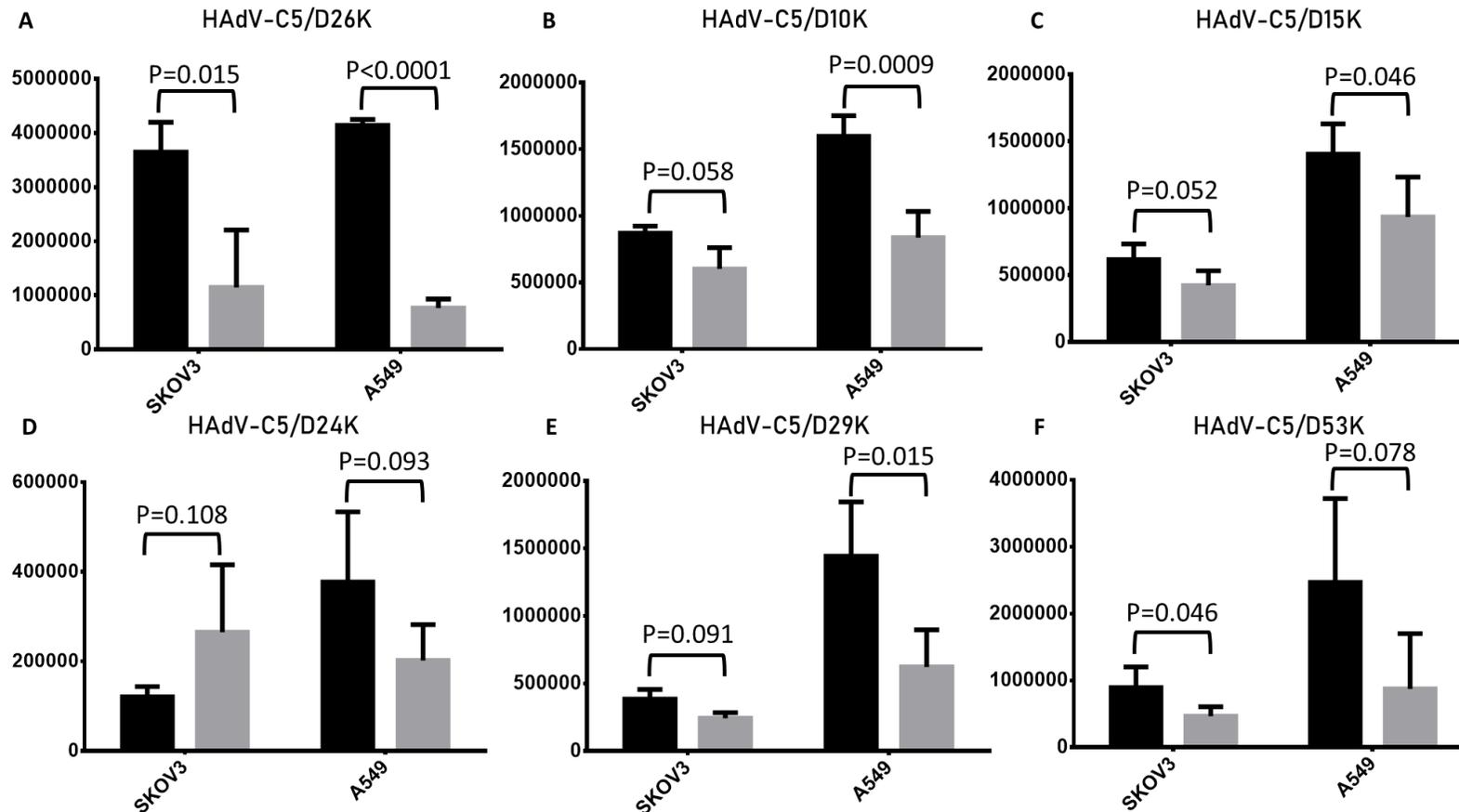


Figure 5-16 Sialic acid cleavage significantly reduces the infectivity of 5 of 6 tested HAAdV-C5 viruses pseudotyped with species D fiber-knob domains. After neuraminidase treatment the infectivity of HAAdV-C5/D26K was significantly inhibited in SKOV-3 and A549 cells (A). HAAdV-C5/D10K saw a non-significant reduction in infectivity in SKOV-3 cells, but was significantly inhibited from infecting A549 cells (B), the same is true of HAAdV-C5/D15K infectivity (C). HAAdV-C5/D24K was not significantly inhibited from infecting either cell line (D), while HAAdV-C5/D29K was significantly inhibited from infecting A549 cells, but only a non-significant reduction was observed in SKOV-3 cells (E). Conversely the HAAdV-C5/D53K pseudotype was significantly inhibited from infecting SKOV-3 cells but not A549 cells where a non-significant decrease was seen (F). Cells were infected with replication incompetent virus expressing a luciferase transgene. $n=4$, Error= \pm SD.

5.7 Discussion

In this study we sought to evaluate whether HAdV-D26 could utilise sialic acid as a cell entry receptor. We demonstrated that HAdV-D26K can form a stable complex with sialic acid through a combination of crystallographic and biological assays. We observed that it is likely that other members of the species D adenoviruses bind sialic acid. We investigated this further and observed that HAdV-D29K can also form a complex with sialic acid, and that HAdV-D30K appears to bind to a small molecule ligand as well, though at this stage we cannot conclusively define the identity of this ligand. Neuraminidase assays confirm sialic acid usage by HAdV-D29K in A549 cells, and significant reductions in infectivity after neuraminidase treatment for several other species D adenoviruses in SKOV-3 and/or A549 cells. Taken together this suggests that sialic acid may be used widely across the species D adenoviruses and identified a broadly conserved sialic acid binding pocket.

We demonstrated that HAdV-D26K can form a stable interaction with sialic acid (Fig5-4), and that this interaction can be used to initiate infection (Fig-3). Cleavage of sialic acid by neuraminidase treatment was seen to reduce the infectivity of the HAdV-C5/D26K pseudotyped virus. After neuraminidase treatment the total fluorescence intensity of cells infected by the HAdV-C5/D26K virus were reduced to levels slightly above those seen in the comparable cell lines infected by HAdV-C5, which infects via CAR (Fig5-3B,C). It is plausible that the residual levels of infection seen post-neuraminidase treatment in HAdV-C5/D26K are due to interaction with CAR, which is expressed only at low levels in these cells, and which HAdV-C5/D26K was seen to be able to bind to with weak affinity⁷², and residual sialic acid not removed during the neuraminidase treatment. Future work could investigate this by simultaneously blocking CAR and sialic acid binding.

5.7.1 pH is unlikely to affect sialic acid binding to HAdV-D26K

We were able to solve the structure of sialic acid in complex with HAdV-D26K at pH4.0 (Fig5-4.A,B,C) and pH8.0 (Fig5-4D,E,F). The differences observed in the interactions at each pH were in the conformation of the glycerol arm, an additional water bridge at pH8.0, and double conformer of Lys349 (Fig5-6). The Lys349 double conformer and additional water bridge are likely due to

differences in resolution. Lysine has a long side chain and is likely to be at least partially flexible, the high resolution of the data obtained allows both the preferred conformations to be seen in the density, giving the double conformer. Since both the observed conformations maintain the polar contact with sialic acid it is unlikely to affect the sialic acid binding. Similarly, the additional water observed in the pocket for the pH8.0 structure of sialic acid in complex with HAdV-D26K is likely due to the slightly increase resolution and/or averaging of the data, rather than a biologically important difference between the two pH's. However, we cannot rule out that the difference in electrostatic potential observed when the fiber knob protein is at different pH's could alter the ability to form polar contacts and coordinate water molecules.

5.7.2 HAdV-D26K contains residue substitutions in the sialic acid binding pocket enabling more direct bonds

Examination of the intensity of the electron density around sialic acid in complex with HAdV-D26K at each pH shows the possible positions of the arm (Fig5-7A,B). In comparison to HAdV-D37K²⁷ and D29K (Fig5-14A,B), HAdV-D26K has a well localised glycerol arm. However, more than one conformation is visible at both pH's (Fig5-6,7). This appears to be because the long Gln348 side chain and the orientation of Thr310 side chain (Fig5-8) means that regardless of the orientation of the glycerol arm it can still form similar polar contacts and water-bridges (Fig5-6). The high number of glycerol arm conformations observed is therefore likely to be due to the plasticity in the interface enabling multiple conformations, and the high resolution of the dataset enabling even rare conformations to be observed in the density.

The side chains which enable HAdV-D26K to form direct interactions with the sialic acid glycerol arm are an important difference with the best studied sialic acid binding adenovirus, HAdV-D37K^{27,147}. HAdV-D37K also forms direct polar contacts with Lys345 and Tyr312 (Fig5-8B) in the same manner as HAdV-D26K has equivalent contacts to Lys349 and Tyr314 (Fig5-8A), as well as sharing a main chain oxygen contact with the sialic acid nitrogen. However, the rest of the HAdV-D37K bonds to sialic acid are water bridges due to the short side chains of Thr310 and Ser344. In contrast HAdV-D26K possesses long side chains at the equivalent Asn312 and Gln348 residues, respectively, as well as a polar

Thr319 rather than Pro317. This residue substitutions enable direct polar contacts, which are observed to be complemented by water bridges, due to the closer proximity of the charge residues side chains to sialic acid. This results in HAdV-D26K having a much more complex and extensive network of sialic acid binding interactions compared to HAdV-D37K.

This closer sialic acid binding contact in HAdV-D26K compared to HAdV-D37K also extends to the hydrophobic interface observed with the N-acetyl methyl group. In HAdV-D37K the hydrophobic interface is formed by Tyr308 at the base with Tyr312 and Val322 forming hydrophobic surfaces lateral to the methyl group of sialic acid (Fig5-8D). In HAdV-D26K the hydrophobic interface is formed by Ile310 at the base, with Tyr310 and Ile324 laterally (Fig5-8C). The HAdV-D26K hydrophobic interface uses the long and short arms of the Ile324 side chains, respectively, to form a cradle, with the long Ile324 arm remaining mobile (Fig5-7C,D) and creating the correct space for the methyl group next to the Tyr314 hydrophobic face. This results in a much closer hydrophobic interface leaving far less space for the methyl group to move in, and correctly orienting the opposing N-acetyl oxygen to form contacts with Asn312.

5.7.3 Regions of weak density around residues with long side chains may indicate an induced fit model of sialic acid binding

The mobile Ile324 arm in HAdV-D26K is likely to be an example of induced fit (Fig5-7C,D). Similarly, the Gln348 residue is seen to have two possible conformation in both pH4.0 and pH8.0 crystallisation conditions. Conformation A (Fig5-7A,B) forms polar contacts with the glycerol arm of sialic acid whereas the B conformation hangs in space. In both cases we cannot determine whether the binding of sialic acid results in the conformation observed with sialic acid bound, or whether that conformation is induced upon proper collision with sialic acid. The most plausible explanation would seem to be that Gln348 is mobile and, upon sialic acid binding, prefers the conformation in which it contacts the glycerol arm as it is lower energy, “locking” the arm in position after binding, regardless of the arms conformation upon the Gln348 forming the bond. While the Ile348 long arm would seem likely to be pressured into this conformation as sialic acid collides and forms bonds elsewhere. However, we cannot rule out other induced fit mechanisms without further studies.

The sialic acid binding pocket in HAdV-D26K can, therefore, be summarised as three synchronous mechanisms. First, an N-acetyl anchor, in which the oxygen forms both a direct polar contact and a water bridge to Asn312 and the methyl group is held in position by a hydrophobic cradle. Second, an inducible lock in which Gln348 bonds the glycerol arm in place. Third, a network of polar contacts to the carboxyl group, C4-OH, and nitrogen atoms around the central pyranose ring.

5.7.4 HAdV-D26K and HAdV-D19pK may belong to a wider subset of non-EKC causing sialic acid binding adenoviruses

Taken together thus shows HAdV-D26 to have a much more sophisticated mechanism of sialic acid interaction than HAdV-D19p and the EKC causing adenoviruses, suggesting HAdV-D26 is likely to have a higher affinity for sialic acid.

This was an unexpected result as HAdV-D26K is dissimilar to the known sialic acid binding species D adenovirus fiber knobs. Sequence analysis of the HAdV-D26 fiber knob amino acid sequence highlighted that while many residues which are known to contact sialic acid in other fiber knob proteins are conserved (Fig-1A) the overall sequence is dissimilar, with just 56.76% similarity to the HAdV-D37K sequence (Fig5-1B). The electrostatic surface of HAdV-D26K is more acidic than any of the previously identified species D sialic acid binding fiber knobs (Fig5-1B, Fig5-2D). The fiber knobs surface electrostatic profile showed some degree of similarity with that of HAdV-D19pK (Fig5-2). While HAdV-D26 has been implicated in EKC³⁴⁵ the instances are very rare and it is generally associated with respiratory disease^{393,394}. HAdV-D19p prototype was discovered in an ocular conjunctival infection, but has not been isolated in humans since its initial discovery in Saudi Arabia in 1955, and has never been associated with EKC³⁹⁵⁻³⁹⁷.

Alignment of all species D fiber knob sequences shows that many of the HAdV-D26K residues which contact sialic acid are conserved in most or all the other viruses (Fig5-9). For example, tyrosine which contacts C4-OH is completely conserved, as is the Lys349. Those residues which are not conserved are often functionally homologous, such as Ile310 in HAdV-D26K is valine in most other species, while the equivalent residue in HAdV-D37K is tyrosine, other species

have a phenylalanine. All those options are possible of fulfilling the role as the base of the hydrophobic interface.

Having determined homology between sialic acid contact residues at a sequence level in the species D adenoviruses we wished to investigate the structural homology. We have solved the crystal structures of seven previously undetermined species D adenovirus fiber knob proteins belonging to HAdV-D10, D15, D26, D29, D30, D48, and D49 (see previous chapters, Fig5-5, Fig5-10 to 13). Using these structures, we visualised the known sialic acid binding residues of HAdV-D26K and HAdV-D37K and compared them to the homologous residues of the other virus fiber knob proteins (Fig5-15).

We saw broad conservation of sialic acid contact residues in species D fiber knob proteins commensurate with what would be expected from the sequence analysis (Fig-9, Fig5-15). An interesting finding is that there appear to be two variations seen in the hydrophobic pocket. HAdV-D37K is prototypical of the first type, in which a hydrophobic ring structure (tyrosine) forms the base of the hydrophobic patch. This creates a flat hydrophobic patch in which the sialic acid methyl group may retain flexibility (Fig5-8D, Fig5-15). This is also seen in HAdV-D10K and D29K (Fig5-15). The second type is that exhibited by HAdV-D26K in which hydrophobic side chains make up the base of the hydrophobic interface forming a “cradle” around the N-acetyl methyl group (Fig5-8C, Fig5-15). This is exhibited in HAdV-D15K, D30K, D48K, D49K. Sequence alignment data shows that all residues positioned at the base of the hydrophobic patch conform to one of these two models (Fig5-9). Either they are tyrosine or phenylalanine, suggesting the flatter interface, alternatively they are valine or isoleucine suggesting the cradle interface.

5.7.5 Fiber knob binding to sialic acid does not preclude the possibility of binding to alternative sugar monomers

Recent work has shown that HAdV-D37K can form interactions with sulphated glycosaminoglycans without initiating infection, acting as “decoy receptors”³⁶⁴ and that that these decoy interactions may make effective drug targets³⁶⁵. Interactions were detected with glycans terminating in residues other than sialic acid, including mannose, glucuronic acid, unsaturated hexaronic acid, and N-

acetyl glucosamine. This shows that interaction with glycans is not necessarily sialic acid dependent.

Given this information that the sialic acid binding interface is conserved, but species D fiber knob glycan binding may not be limited to sialic acid we investigated the possibility of interaction with sugar monomers other than sialic acid. To this end we performed soaking experiments on species D fiber knob crystals using equimolar concentrations of 5 different glycan terminating sugar residues: galactose, mannose, fucose, N-acetyl glucosamine, and sialic acid (Neu5Ac). We theorised that we would see density for the highest affinity interaction, given high enough resolution diffraction.

Soaking of HAdV-D29K crystals showed clear density for sialic acid, having both a glycerol arm and an N-acetyl group (Fig5-14A,B). In one of the monomer copies the sialic acid was relaxed (Fig5-14A) forming a similar confirmation to that seen in HAdV-D26K (Fig5-4C,F), but in the other the pyranose ring is forced into a boat conformation to accommodate the glycerol arm being angled away from the three fold axis (Fig5-14B), creating strain. This unlikely pose may be a result of error in the data refinement as the overall sialic acid density for the glycerol arm was weak. Therefore, repeating the crystallisation with a longer sialic acid soak may enhance the quality of the density and allow better definition of the sialic acid binding location to be achieved.

Sialic acid is the largest of the 5 molecules which were soaked into the HAdV-D29K crystals. Since all crystallographic data represents an average of all the repeating units within the crystal it is possible we are seeing the outline of sialic acid but in some monomers different sugar monomers are interacting. Due to this we cannot preclude the possibility that other residues in the soaking mixture may interact at this site. Further soaking experiments will need to be performed in the absence of sialic acid to assess this.

Sugar soaked HAdV-D30K crystals also showed density for a ligand partner in the same location as has been observed for HAdV-D19pK, D37K, D26K, and D29K. However, the data was relatively low resolution. Refinement was attempted with all 5 possible ligands, and the resultant maps all provided plausible explanations for the density. With this data set it is not possible to

identify which of the 5 sugar monomers which were soaked into the crystal is the binding partner.

HAdV-D15K did not show any clear density for ligand binding upon soaking (Fig5-15A,B). This can be explained by 2 factors. First, the close packing arrangement does not afford easy diffusion of the small molecules through the crystal, limiting the effectiveness of the soak (Fig5-15C,D). The soaking experiment supports this, as the non-soaked crystal structure crystallised in $P 2_1 2_1 2_1$, while the soaking broke symmetry to $P 1$. Longer soaks were attempted but resulted in the crystals dissolving. This suggests that sialic acid interaction disrupts the crystal packing arrangement.

The second reason explaining the lack of sialic acid observation during soaking experiments with HAdV-D15K is seen in the sialic acid binding pocket observed in Figure 15. HAdV-D15K is unique among the investigated fiber knob proteins in having an isoleucine residue at position 319. In the conformation observed in the unliganded crystal structure of the predicted sialic acid binding residues the long arm of this residue is extending into the hydrophobic interface where it would inhibit the N-acetyl methyl group from binding (Fig5-13). In the soaked crystal the structure the pocket in which some density can be observed in the sialic acid binding this isoleucine is oriented with the long isoleucine arm away from the hydrophobic interface, potentially make space for the N-acetyl methyl group (Fig5-15A). Conversely, the isoleucine is oriented in the inhibitory conformation in the pocket when negative density was observed (Fig5-15B). Taken together, this data regarding the HAV-D15K interaction with sialic acid suggests that the sialic acid may be incapable of binding due to crystal properties rather than biological ones. Co-crystallisation of HAdV-D15K may afford the greater flexibility for the isoleucine residue and enable sialic acid binding, if the sialic acid does indeed interact.

5.7.6 Neuraminidase assays suggest many species D adenovirus fiber knobs can utilise sialic acid as a receptor to varying degrees

Neuraminidase assay data suggest the possibility that HAdV-D15K, and other tested species D adenovirus fiber knob proteins, can bind to sialic acid and utilise it to form an infection (Fig5-16). This assay found significant reduction in the infectivity of HAdV-C5 pseudotyped with species D fiber knobs for all the

tested viruses, except for HAdV-C5/D24K. However, the reduction was not significant in all cell lines tested. This assay utilised luciferase expressing adenovirus pseudotypes, necessitating the use of a luminescence read out and controlling the experiment by protein concentration in the well, measured by BCA assay. Therefore, there are numerous points in this assay where error could be introduced, such as warming of the cells after neuraminidase cleavage resulting in regeneration of the sialic acid bearing glycans, inaccurate measurement of protein concentration in the well, protein concentration not being representative of luciferase concentration. Generation of GFP expressing viruses would eliminate much of this experimental error and enable more accurate determination of the effect of cell surface sialic acid on the transduction ability of these pseudotypes. Until such a time as this data can be replicated with improved resolution no definite statements can be made regarding the sialic acid binding status for these viruses.

Despite this uncertainty, it seems likely that many other species D adenovirus fiber knobs may be capable of binding to sialic acid. This sialic acid binding seems to be capable of forming a productive infection in the context of the HAdV-C5/D26K pseudotype, making it seem likely that the native viruses also display some degree of affinity for the sialic acid receptor. Indeed, HAdV-D37, D53, D54, D64 (formerly HAdV-D19a), and D8 (identical to that of HAdV-D54) have been shown to bind to sialic acid in the context of the wildtype virus and initiate infection. However, the degree of CAR binding in most species D adenoviruses remains uncharacterised as well, therefore it is equally plausible that many will utilise CAR over sialic acid.

HAdV-D37K has been shown to have preferential binding to the di-sialylated GD1a glycan¹⁴⁷. However, recent data shows that this interaction, whilst apparently preferential, is not specific³⁶⁴. Given the apparent plasticity in the interface and the high degree of structural similarity between sugar monomers, it would be more surprising if the interaction was entirely sialic acid specific. If HAdV-D37K has this flexibility in the types of glycan chains which it can interact with it seems likely that it will transfer to other fiber knob proteins which bind to sialic acid in similar way, which seems to include much if not all of the species D adenoviruses.

The various fiber knob proteins display variable affinities for sialic acid and have very different isoelectric points. For example, HAdV-D37K and HAdV-D19pK, while HAdV-D37 is strongly associated with EKC HAdV-D19p is not. Similarly, HAdV-D26 is more generally associated with ARDS than with EKC. HAdV-D26 and D19p have more acidic isoelectric points than HAdV-D37K and the rest of the EKC associated adenovirus fiber knobs (Fig5-2)^{27,59}.

Given that these fiber knobs bind to sialic acid but are not all associated with ocular infections they must be utilising sialic acid to infect other tissues. It therefore seems logical that the sialic acid binding adenoviruses have preferences to different glycans. The availability of suitable glycan binding partners may be the means which controls tissue tropism. Glycan arrays, as performed by the Arnberg lab in the identification of GD1a as a binding partner, may present an effective means of determining specifically

Presently, the investigation of sialic acid binding adenoviruses is largely focused upon the EKC causing adenoviruses. As a result, most papers use either A549 (lung carcinoma) cells or human corneal epithelium (HCE) cells as a model of infection in biologically relevant tissue. A potential pitfall of this approach is that, while it is an effective model under these specific circumstances, different cell lines have very different glycosylation patterns. Therefore, ruling out viral infectivity in one or two cell lines does not exclude infection in that tissue. A better alternative would be to match the results of virus binding to glycan micro-array vs the glycosylation profiles of various tissues.

These findings build upon the increasingly complex body of knowledge surrounding the species D adenoviruses which account for nearly 70% of the known adenovirus diversity, both in terms of the fiber knob and total species diversity⁷². The knowledge that these viruses utilise sialic acid to infect cells raises the prospect of using antiviral drugs which mimic it, including the ME0462 compound being developed out of the Arnberg lab⁵⁹, suramin³⁶⁵ which is already approved for the treatment of African sleeping sickness, and the host of neuraminidase inhibitors developed against influenza such as Zanamivir¹⁴⁸.

5.8 Summary of chapter 5

The major findings in this chapter are as follows:

1. HAdV-D26K binds to sialic acid and HAdV-C5/D26K can utilise sialic acid as a primary cell entry receptor, as shown by the inability to infect cells efficiently in the absence of cell surface sialic acid (Fig5-3).
2. The mode of sialic acid interaction is like that of EKC causing adenoviruses as determined by comparison of new HAdV-D26K crystal structures in complex with sialic acid (Fig5-4) with HAdV-D37K in complex with sialic acid (Fig5-6 to Fig5-8).
 - a. This suggests that sialic acid binding alone is not the driver of EKC pathogenesis, as if it were then HAdV-D26 would be associated with EKC as well.
3. The residues involved in sialic acid binding in HAdV-D26K and the EKC causing adenovirus fiber knob proteins are widely conserved in the species D fiber knob proteins (Fig5-9 and Fig5-13).
4. Sialic acid is likely to be a receptor for other species D adenoviruses as inferred by analysis of figures 5-9 to 5-15.

CHAPTER 6. General Discussion

The focus of this study has been to examine the diversity present in the species D adenovirus fiber knob proteins. This is undertaken with a view to how this influences Adenovirus receptor usage, and how this knowledge can be applied to the engineering of improved adenovirus derived therapeutic vectors.

This work reveals an exceptionally broad tropism mediated by the HAdV-D49 fiber knob protein in chapter 3. HAdV-D49 had previously been demonstrated to efficiently transduce vascular smooth muscle and endothelial cells³⁵⁰. In chapter 4 we describe how three species D adenovirus fiber knob proteins, belonging to HAdV-D26, HAdV-D48, and HAdV-D49, do not appear to utilise CD46 as a primary cell entry receptor as had previously been reported for all three of these viruses^{41,103,351,381}. Instead, we show that these three viruses all have a limited CAR tropism. While the affinity of this CAR tropism is variable, likely moderated by steric inhibition caused by variable DG loop sequences, they all bind with significantly lower affinity than HAdV-C5, suggesting that CAR may be a receptor of last resort, rather than the key tropism determinant. In chapter 5 we identify what appears to be the primary receptor for HAdV-D26 fiber knob protein as sialic acid, a receptor which has only previously been shown to be utilised by EKC causing adenoviruses. We conclude by showing broad conservation of key sialic acid binding residues across the species D adenovirus fiber knobs and show significant decreases in transduction affinity of pseudotypes bearing species D fiber knob proteins after treatment with neuraminidase to remove cell surface sialic acid.

The discovery of a broadly occurring adenovirus tropism, as was observed for HAdV-D49K, is not in itself surprising. Previously identified receptors for adenoviruses are also broadly expressed. CD46 is found on the surface of all nucleated human cells³⁹⁸⁻⁴⁰⁰, whilst CAR is expressed on the surface of platelets, erythrocytes, and most epithelial tissues, and is especially highly expressed in gastrointestinal tissues and the skin^{401,402}. Desmoglein 2 has a similar expression profile to CAR, highly expressed on epithelial tissues and gastrointestinal tissues, likely due to its similar function in intercellular adhesion^{403,404}. Sialic acid is similarly broadly expressed, as it is the terminal residue of complex glycosylation motifs found on most human cells, though the distribution of sialic acid is unlikely to correlate exactly with receptor interaction due to the influence the variable glycan chains supporting the sialic acids⁴⁰⁵.

While we were able to exclude previously described adenovirus receptors CD46, DSG2, and sialic acid as primary cell surface receptors for HAdV-D49K, we were unable to exclude a mechanistic basis for this broad tropism derived from its CAR binding ability or potential non-sialic acid based glycan interactions.

6.1 Adenovirus receptor tropism as a driver of viral evolution

In evolutionary terms it is preferable for a viral cell receptor to be broadly expressed, so long as the cells it is expressed on are permissive to replication. In theory the more the cell entry receptor for a virus is expressed the more opportunities it has to infect cells and replicate, resulting in an evolutionary advantage. On a population level this is advantageous, though is not beneficial to the individual viral genome as it can lead to super-infection by another strain of the same virus with the potential to dilute the pool of progeny and outcompete it. This behaviour is observed in numerous viruses, including hepatitis C, Influenza A, and phage viruses, among others^{406–410}, but not in adenoviruses. Indeed, adenoviruses exhibit the inverse, where superinfection enables recombination between virus strains and the generation of new viral serotypes, representing a more cooperative mode of viral evolution^{21,32,376,411,412}. This often leads to the “swapping” of proteins, for example the recently (2019) described isolates HAdV-D100, HAdV-D102, and HAdV-D103, all possess identical fiber proteins to HAdV-D30, but with hexon proteins belonging to HAdV-D17, HAdV-D38, and HAdV-D33, respectively, and newly described penton proteins^{25,413–415}.

One inference from our phylogenetic comparison of adenovirus genome variation to fiber knob amino acid sequences (Fig3-1A, B) is that different adenovirus species may have differing evolutionary strategies in terms of receptor utilisation. Species D fiber knob receptors exhibit more diversity, individually, than the average for their genomes. Species B fiber knob proteins exhibit the opposite trend, with less diversity in the fiber knob sequence compared to their average genome diversity. The species B fiber knobs all use either DSG2, CD46, or both as primary cell entry receptors, as far as is known^{55,89,370,371}. Our previously published work, and the data in this thesis (chapters 4 and 5), suggest that many of the species D adenoviruses could

utilise some combination of CAR and sialylated glycans as their primary receptors^{72,356}.

It therefore seems plausible that the species B adenoviruses employ a strategy reliant upon near universal cell infection, mediated by their ubiquitous primary cell receptors. They then have more specialised post-cell entry replication machineries to allow efficient replication in tissues for which it is adapted. In the case of HAdV-B3 and HAdV-B7 this is epithelial airway tissue, leading to acute respiratory disease syndrome (ARDS), sometimes even in healthy adults^{416–418}. Therefore, adaptations of these viruses for therapeutic use are more likely to generate specialised post-cell entry machineries, as is observed in Enadenotucirev, than adaptations of the capsid and cell entry machinery^{102,291}. Genetically, this would result in a convergence of the sequences of the capsid into a few clades, while the rest of the genomes diverged, as we witness in the analysis in Figure 3-1.

The inverse strategy may be apparent in the species D adenoviruses. A wide variety of cellular host receptors and a less specialised replication machinery, leading to diverse capsid proteins, with reduced diversity in the rest of the genome.

Sialic acid as a primary cell entry receptor presents a challenge to any virus using it. Sialic acid is only the terminal residue of a longer glycosylation chain, the exact sequence of which is variable depending upon a variety of conditions, such as tissue type, environment, and available protein carriers, creating an ever shifting glycosylation landscape^{419–421}. This requires the pathogen to adapt to use types of sialylated glycan with amenably charged chains. This drives enormous diversity in the sialic acid binding receptor protein, requiring rapid adaptation from any pathogen utilising these sialylated glycan chains as entry receptors. The best studied example of this is influenza. Influenza has seasonal variants, which are caused by recombination events between flu viruses which cause novel combinations of, and recombination between, its receptor proteins, neuraminidase and haemagglutinin^{149,422–424}.

Were species D adenoviruses using sialic acid as a broadly applicable receptor we would expect to see three indications:

1. Large amounts of diversity in the receptor proteins, enabling binding to diverse sialylated glycan chains.
2. Low levels of pre-existing seroprevalence as a result of the shifting viral capsid morphology.
3. Varying, but not strict, sites of pathogenesis due to the use of different specific glycan moieties, preferentially expressed on different tissues.

This is exactly what is observed in the species D adenoviruses. Species D adenoviruses account for more than 50% of total adenovirus diversity, most of which occurs in the capsid^{20,32,355,425}. The human adenovirus working group isolates and defines new adenovirus serotypes strictly based upon their capsid proteins. The vast majority of new adenovirus types are species D, and the most commonly varied capsid component is the fiber protein, there are currently 84 different fiber proteins defined²⁵. This fulfils the first criteria.

To fulfil the second criteria, we look to the previously discussed information about adenovirus seroprevalence. Species D adenoviruses are rarely occurring, and have low rates of seroprevalence in the population as determined by the prevalence of neutralising antibodies^{37,39,340,350,351,426}. Though it appears certain groups may have higher rates of nAb activity against particular serotypes than the rest of the population. For example men who have sex with men were found to have higher rates of HAdV-D26 nAbs than the general population⁴²⁷.

nAbs are not only found against the fiber protein, which would be most selected against in this mode of evolution, but also other capsid proteins, especially the hexon. Previous studies have suggested that the hexon is the primary target of nAbs²⁰⁵. However, follow-up studies find the fiber can be a more important target for nAbs during natural adenovirus infections, as opposed to deliberate administration. Patients are observed to be capable of neutralising adenoviruses based only on anti-fiber nAbs²⁰⁶. Clearly nAb activity is not limited to any one adenovirus capsid protein, making the fiber only a part of the picture in terms of anti-adenovirus immunity. However, any fully neutralising antibody directed against the fiber would, conceivably, be enough to create a selective pressure which could be evaded by rapidly shifting fiber morphology.

Finally we suggest that any sialic acid using species of virus would have broad tissue tropism and varying sites of pathogenesis due to preference for different

glycopatterns, as observed in influenza^{428–430}. Indeed, species D adenoviruses are most commonly associated with epidemic keratoconjunctivitis^{19,365,377,388}. However, they are also associated with acute respiratory illness^{340,418,431} and gastroenteritis^{138,340,427,432}, as well as from the faeces of asymptomatic adults³⁴⁶ and HIV patients^{347,433}. A study of clinical symptoms associated with species D adenovirus infection causing diarrhoea, nausea, abdominal pain, ocular disease, and respiratory disease³⁴⁰. HAdV-D26 alone has been associated with EKC, ARDS, HIV, and healthy patient faecal matter, the latter being the original isolation^{345,377,393,434,435}. This indicates species D adenovirus infection of three sites, the eye, gastrointestinal tract, and respiratory tract, all heavily sialylated and exposed mucus membranes.

6.2 Further investigation of species D adenovirus receptor tropism is imperative to understanding its pathogenesis

While the evidence in this thesis, previous publications, or the discussion above absolutely do not conclusively show that all species D adenoviruses utilise sialic acid as a primary cellular receptor, it does provide a plausible explanation for behaviour which is seemingly unique to species D, within the adenoviruses. It also presents avenues of investigation to begin unpicking the species D adenoviruses phenotypes. First, more extensive phylogenetic comparisons of the species D and B adenoviruses can help define the genes which contribute more and less to the species genetic diversity. Second, neuraminidase assays such as the ones performed in this thesis can be utilised to investigate more species D adenovirus species, both as fiber knob pseudotypes and in the whole virus species.

It is also crucial to understand the role of CAR in species D adenovirus biology. Our data demonstrates a spectrum of different CAR binding affinities by different species D adenovirus fiber knob proteins (chapter 4)⁷². Whilst the utilisation of sialic acid bearing glycans as primary receptors does explain much of the activity observed in the species D adenoviruses, it fails to address the role of CAR. It is possible that CAR truly is an incidental binding activity, an evolutionary vestige, but more likely it is a viable tropism in the absence of a preferable glycan. It thus seems conceivable that the species D adenoviruses

exist on a spectrum between largely CAR dependent to largely sialylated glycan dependent, though the nature and breadth of this distribution is unknown.

Regardless of the receptor tropism of the entire species, evidence in this thesis, and our recent manuscript³⁵⁶, provides the first examples of non-EKC associated sialic acid utilising species D human adenovirus fiber knob proteins. The only other known example of a non-EKC associated sialic acid binding human adenovirus is HAdV-G52, which is reported to utilise polysialic acid as a receptor⁵⁷. HAdV-G52 has been isolated only once, from a patient with gastroenteritis, and possesses two different fiber proteins (long and short). The short fiber protein binds to sialic acid, while the long binds to CAR, granting it the same overall cell binding characteristics as we propose for the species D adenoviruses and observe in HAdV-D26K, but by a different means¹³⁸.

Therefore, any complete understanding of species D adenovirus tropism depends on understanding of their interactions with sialylated glycans. Recent work from the Arnberg lab shows that non-sialylated glycans may also play a role as decoy receptors^{59,364}. The Arnberg lab also previously showed that HAdV-D37 had a preference for GD1a glycan, which helps drive its ocular tropism, by use of glycan array¹⁴⁷. Exploration of adenovirus diversity by this method would reveal any glycan preferences, so long as the glycans of interest are present on the chip. A less biased method may be the use of cross-linking mass-spectrometry to pull down cell associated glycans. However this technology is still emerging^{436,437}.

6.3 The potential for species D adenoviruses as therapeutics

The findings in this thesis demonstrates that species D adenoviruses an even more attractive therapeutic platform than previously believed. The previously discussed low levels of pre-existing immunity in the population make species D adenoviruses less likely to be rapidly cleared by the human immune system, and therefore more likely to achieve therapeutic effect. The diversity observed in receptor tropism opens several intriguing possibilities, both as retargeted oncolytic adenoviruses, and in other modes of virotherapy.

6.3.1 Towards target specific species D adenovirus vectors

It appears that some degree of CAR utilisation as a cell surface receptor may be retained in HAdV-D26, D48, and D49, as observed in chapter 4. Many of the key CAR binding residues are conserved across the species D adenovirus, implying that these are unlikely to be isolated incidences. There are previously described mutations in HAdV-C5 which can ablate CAR binding affinity, such as the KO1 and TAYT mutations^{74,76}. By applying the mutations to other adenovirus fiber knobs it should be possible to ablate CAR binding.

Similarly, knowledge of the sialic acid binding domain of HAdV-D26, and other species D adenoviruses enables engineering of ablation mutations. Whilst no sialic acid binding mutation have previously been described, reversing the nature of the key sialic acid binding residues presents clear candidates for sialic acid binding ablation mutations. Mutation of the tyrosine residue which binds to the C4 oxygen of sialic acid (Tyr312, Fig5-8) is likely to be particularly effective as it would disrupt both the hydrophobic interface and remove the most completely conserved polar contact (Fig5-9). It is clear, however, that there is a high degree of plasticity in the sialic acid binding motif and that much of it may derive from the electrostatic surface potential. As such any mutation to ablate sialic acid binding should be conscious of this and aim to create a more negatively charged surface to repel sialylated glycans entirely, rather than rely entirely upon disruption of the anchor.

There has already been a described attempt to generate a targeted HAdV-C5/D48K virus, with a view towards developing an oncolytic virotherapy, by the Parker group⁷⁹. This involved the insertion of an $\alpha\beta 6$ specific A20 peptide into the DG loop of the HAdV-D48 fiber knob (HAdV-C5/D48K.DG-A20). In $\alpha\beta 6$ expressing BT-20 cells both variants were effective at enabling expression, while the unmodified HAdV-C5/D48K pseudotype exhibited only basal infectivity. However, in patient derived ovarian ascites cells (EOC) these viruses were outperformed by a KO1 (CAR binding ablated) mutated HAdV-C5 virus expressing the A20 peptide in the HI loop.

Interestingly the HAdV-C5/D48K.DG-A20 outperformed the HAdV-C5/D48K control by 10-fold in CAR expressing EOC. However, in low CAR EOC the HAdV-D48K.DG-A20 outperformed the control by 66-fold. Given the new

information regarding HAdV-D48K CAR binding presented in chapter 4 it seems likely that the reason for this difference was enhanced transduction by the control HAdV-C5/D48K in CAR expressing cells, making the improvement by the virus expressing A20 in the CAR inhibitory DG loop (Fig4-12) relatively weaker.

Given this new information regarding HAdV-D48K receptor tropisms it would be extremely interesting to revisit the $\alpha\beta 6$ targeted HAdV-C5/D48K pseudotype. The addition of CAR ablation mutations may reduce off target infection. When compounded with sialic acid ablation mutations, should HAdV-D48K have such an interaction, this could yield a truly specific vector, with low levels of pre-existing immunity in the population. Success in this pseudotype would warrant exploration of this in the whole HAdV-D48 type. The same approach could be taken to the design of similar HAdV-D26 based vectors, where the effectiveness of sialic acid binding ablation mutations could be more easily assessed given its clear sialic acid binding ability (Fig5-3)³⁵⁶.

6.3.2 The potential of species D adenoviruses in vaccine development

HAdV-D26 has seen significant development as a vaccine, as has been previously discussed (chapter 5)^{40,105,369}. HAdV-D49 has also been subject to interest as a vaccine vector³⁵¹. Whilst the Ad26 vaccine has been highly effective the mechanisms which make it so capable of generating a lasting immune response remain unclear.

The discovery that HAdV-D26 can utilise sialic acid as a cell entry receptor may provide a new insight into what make this virus so immunostimulatory. Binding to cell surface sialic acids will necessarily desialylate the cell surface as the virus becomes internalised, taking the sialic acid bearing glycan with it. This has previously been observed in influenza⁴³⁸. Desialylation of tissues has been shown to promote inflammation^{439,440}, promoting cytokine release by monocytes⁴⁴¹, and even stimulating the recruitment of CD8⁺ T-cells to the site of desialylation⁴⁴². It would therefore seem logical that a virus such as HAdV-D26 which utilises sialic acid as a receptor could help stimulate greater immunogenicity than a non-sialic acid binding virus. This would assist the formation of lasting immunity by the recruitment of memory forming CD8⁺ T-cells

to the area in which viral antigens, or antigens added as a transgene in a vaccine, are being presented after infection.

We also observed in this study that HAdV-D49K enables efficient transduction of dendritic cells (Fig.3-10 to 12). This virus has already been investigated and found to be a potentially useful vaccine vector^{351,443}. The ability to infect DC is particularly appealing, since dendritic cell vaccines have begun to re-emerge as potential cancer treatments, but in the absence of safe and effective tools with which to transduce them^{444,445}.

The field of glycobiology itself is still relatively new, as immunology and virology undergo a renaissance as their roles in cancer immunology and novel therapies become more widely accepted and utilised. As such no robust conclusions can be drawn regarding the role of sialic acid in HAdV-D26's success as a vaccine, nor can any predictions be made about the effectiveness of HAdV-D49 as a DC vaccine. However, both these findings offer intriguing new avenues of exploration for those developing species D adenovirus-based vaccines or studying the role of sialic acid, and other cell surface receptors, in adenovirus pathology.

6.4 Concluding remarks

In this work we have identified the HAdV-C5/D49K pseudotype virus as a potentially useful vector. While the lack of specificity may limit its usefulness as an oncolytic virotherapy its broadly infectious characteristics may hold promise as a gene transfer vector, perhaps for dendritic cell mediated cancer vaccines, fulfilling the first of my aims: to identify potentially useful phenotypes in the species D adenovirus for the treatment of cancer.

We have also clarified that CAR is a receptor utilised by several members of the species D adenoviruses and shown CD46 usage to be unlikely, satisfying the second aim: to clarify species D adenovirus receptor usage. We go on to show that sialic acid is utilised by HAdV-D26K and may be a broadly used species D adenovirus receptor, fulfilling my third aim: to identify previously unknown receptor usage in species D adenoviruses.

The fourth and final aim was to understand the mechanisms which govern species D adenovirus infection. We were successful in this regard in relation to

HAdV-D26, showing sialic acid to be a primary receptor. However, we were unable to clarify the receptor tropism of HAdV-D48K or D49K beyond the observed weak CAR tropism. However, we do demonstrate that species D adenovirus receptor usage may be even more diverse than previously imagined for a virus species of this size. We have begun to illuminate the structures of these viruses receptor enabling future investigation of the tropisms and mechanisms of infection for these viruses.

As with all novel technologies, viral vectors are built upon understanding. Directed evolution experiments and panning of genetic diversity can assist in the development of new agents by presenting candidates with beneficial properties. However, the proper application of these technologies can only be achieved by determining the mechanisms which make them effective. The efforts in this thesis have been towards improving the understanding of how a subset of adenovirus types can infect cells. When adding to the body of knowledge surrounding adenovirus biology we have sought to understand how it impacts development of new adenovirus vectored therapies. The arms race to develop more effective viral vectors for vaccines, oncolytic virotherapy, gene therapy, and beyond, will continue rapidly and is not limited to a single virus type. Therefore, virology must match or exceed this pace in order to enable these new technologies to be properly applied for therapeutic benefit.

References

1. Medawar, P. B. & Medawar, J. S. *Aristotle to Zoos: A Philosophical Dictionary of Biology*. (Harvard University Press, 1985).
2. Guo, Z. S. *et al.* Oncolytic Immunotherapy: Conceptual Evolution, Current Strategies, and Future Perspectives. *Front. Immunol.* **8**, (2017).
3. THE INFLUENCE OF COMPLICATING DISEASES UPON LEUKAEMIA. - ProQuest.
<https://search.proquest.com/openview/eb5cb1ae1913ff84c02cc61afbbd8d5e/1?pq-origsite=gscholar&cbl=41361>.
4. Kelly, E. & Russell, S. J. History of Oncolytic Viruses: Genesis to Genetic Engineering. *Mol. Ther.* **15**, 651–659 (2007).
5. Moore, A. E. The destructive effect of the virus of Russian Far East encephalitis on the transplantable mouse sarcoma 180. *Cancer* **2**, 525–534 (1949).
6. Koprowski, H., Koprowska, I. & Love, R. Ascites Tumor-Virus System as a Biological Tool. *Proc. Natl. Acad. Sci. U. S. A.* **39**, 1147–1148 (1953).
7. Harrington, K. J., Vile, R. G., Melcher, A., Chester, J. & Pandha, H. S. Clinical trials with oncolytic reovirus: Moving beyond phase I into combinations with standard therapeutics. *Cytokine Growth Factor Rev.* **21**, 91–98 (2010).
8. Sanchala, D. S., Bhatt, L. K. & Prabhavalkar, K. S. Oncolytic Herpes Simplex Viral Therapy: A Stride toward Selective Targeting of Cancer Cells. *Front. Pharmacol.* **8**, (2017).
9. Aloni-Grinstein, R., Charni-Natan, M., Solomon, H. & Rotter, V. p53 and the Viral Connection: Back into the Future. *Cancers* **10**, (2018).

10. Haller, O., Kochs, G. & Weber, F. The interferon response circuit: Induction and suppression by pathogenic viruses. *Virology* **344**, 119–130 (2006).
11. Twumasi-Boateng, K., Pettigrew, J. L., Kwok, Y. Y. E., Bell, J. C. & Nelson, B. H. Oncolytic viruses as engineering platforms for combination immunotherapy. *Nat. Rev. Cancer* **1** (2018) doi:10.1038/s41568-018-0009-4.
12. Zygiert, Z. HODGKIN'S DISEASE: REMISSIONS AFTER MEASLES. *The Lancet* **297**, 593 (1971).
13. Hoster, H. A., Zanes, R. P. & Haam, E. von. Studies in Hodgkin's Syndrome: IX. The Association of "Viral" Hepatitis and Hodgkin's Disease (A Preliminary Report). *Cancer Res.* **9**, 473–480 (1949).
14. Baker, A. T., Aguirre-Hernández, C., Halldén, G. & Parker, A. L. Designer Oncolytic Adenovirus: Coming of Age. *Cancers* **10**, 201 (2018).
15. International Committee on Taxonomy of Viruses (ICTV). https://talk.ictvonline.org/ictv-reports/ictv_online_report/.
16. Fields virology - NLM Catalog - NCBI. <https://www.ncbi.nlm.nih.gov/nlmcatalog/101601028>.
17. Jr, S. D. B. *et al.* Adenoviruses Isolated from Saudi Arabia. *Am. J. Trop. Med. Hyg.* **8**, 492–500 (1959).
18. Robinson, C. M. *et al.* Human adenovirus type 19: genomic and bioinformatics analysis of a keratoconjunctivitis isolate. *Virus Res.* **139**, 122–126 (2009).

19. Zhou, X. *et al.* Analysis of human adenovirus type 19 associated with epidemic keratoconjunctivitis and its reclassification as adenovirus type 64. *Invest. Ophthalmol. Vis. Sci.* **53**, 2804–2811 (2012).
20. Robinson, C. M. *et al.* Molecular evolution of human adenoviruses. *Sci. Rep.* **3**, 1812 (2013).
21. Lukashev, A. N., Ivanova, O. E., Ereemeeva, T. P. & Iggo, R. D. Evidence of frequent recombination among human adenoviruses. *J. Gen. Virol.* **89**, 380–388 (2008).
22. Robinson, C. M. *et al.* Computational analysis and identification of an emergent human adenovirus pathogen implicated in a respiratory fatality. *Virology* **409**, 141–147 (2011).
23. Walsh, M. P. *et al.* Computational Analysis Identifies Human Adenovirus Type 55 as a Re-Emergent Acute Respiratory Disease Pathogen. *J. Clin. Microbiol.* **48**, 991–993 (2010).
24. Singh, G. *et al.* Overreliance on the Hexon Gene, Leading to Misclassification of Human Adenoviruses. *J. Virol.* **86**, 4693–4695 (2012).
25. HAdV Working Group, <http://hadvwg.gmu.edu/>. <http://hadvwg.gmu.edu/>.
26. Seto, D., Chodosh, J., Brister, J. R., Jones, M. S. & Community, M. of the A. R. Using the Whole-Genome Sequence To Characterize and Name Human Adenoviruses. *J. Virol.* **85**, 5701–5702 (2011).
27. Burmeister, W. P., Guilligay, D., Cusack, S., Wadell, G. & Arnberg, N. Crystal Structure of Species D Adenovirus Fiber Knobs and Their Sialic Acid Binding Sites. *J. Virol.* **78**, 7727–7736 (2004).

28. Madisch, I., Harste, G., Pommer, H. & Heim, A. Phylogenetic Analysis of the Main Neutralization and Hemagglutination Determinants of All Human Adenovirus Prototypes as a Basis for Molecular Classification and Taxonomy. *J. Virol.* **79**, 15265–15276 (2005).
29. Crawford-Miksza, L. K. & Schnurr, D. P. Quantitative colorimetric microneutralization assay for characterization of adenoviruses. *J. Clin. Microbiol.* **32**, 2331–2334 (1994).
30. Malasig, M. D., Goswami, P. R., Crawford-Miksza, L. K., Schnurr, D. P. & Gray, G. C. Simplified Microneutralization Test for Serotyping Adenovirus Isolates. *J. Clin. Microbiol.* **39**, 2984–2986 (2001).
31. Hierholzer, J. C. & Bingham, P. G. Vero microcultures for adenovirus neutralization tests. *J. Clin. Microbiol.* **7**, 499–506 (1978).
32. Walsh, M. P. *et al.* Evidence of Molecular Evolution Driven by Recombination Events Influencing Tropism in a Novel Human Adenovirus that Causes Epidemic Keratoconjunctivitis. *PLOS ONE* **4**, e5635 (2009).
33. Crawford-Miksza, L. K. & Schnurr, D. P. Adenovirus Serotype Evolution Is Driven by Illegitimate Recombination in the Hypervariable Regions of the Hexon Protein. *Virology* **224**, 357–367 (1996).
34. Espínola, E. E., Barrios, J. C., Russomando, G., Mirazo, S. & Arbiza, J. Computational analysis of a species D human adenovirus provides evidence of a novel virus. *J. Gen. Virol.* **98**, 2810–2820 (2017).
35. Tsai, V. *et al.* Impact of Human Neutralizing Antibodies on Antitumor Efficacy of an Oncolytic Adenovirus in a Murine Model. *Clin. Cancer Res.* **10**, 7199–7206 (2004).

36. Barouch, D. H. *et al.* International Seroepidemiology of Adenovirus Serotypes 5, 26, 35, and 48 in Pediatric and Adult Populations. *Vaccine* **29**, 5203–5209 (2011).
37. Abbink, P. *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).
38. Mennechet, F. J. D. *et al.* A review of 65 years of human adenovirus seroprevalence. *Expert Rev. Vaccines* **18**, 597–613 (2019).
39. Weaver, E. A. & Barry, M. A. Low seroprevalent species D adenovirus vectors as influenza vaccines. *PloS One* **8**, e73313 (2013).
40. Geisbert, T. W. *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).
41. Teigler, J. E., Iampietro, M. J. & Barouch, D. H. Vaccination with Adenovirus Serotypes 35, 26, and 48 Elicits Higher Levels of Innate Cytokine Responses than Adenovirus Serotype 5 in Rhesus Monkeys. *J. Virol.* **86**, 9590–9598 (2012).
42. Crank, M. C. *et al.* Safety and Immunogenicity of a rAd35-EnvA Prototype HIV-1 Vaccine in Combination with rAd5-EnvA in Healthy Adults (VRC 012). *PLOS ONE* **11**, e0166393 (2016).
43. Safety and Immune Response to a Recombinant Adenovirus HIV-1 Vaccine in Healthy Adults - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT00695877>.

44. Tian, X. *et al.* Identification of a Critical and Conformational Neutralizing Epitope in Human Adenovirus Type 4 Hexon. *J. Virol.* **92**, e01643-17 (2018).
45. Choudhry, A., Mathena, J., Albano, J. D., Yacovone, M. & Collins, L. Safety evaluation of adenovirus type 4 and type 7 vaccine live, oral in military recruits. *Vaccine* **34**, 4558–4564 (2016).
46. Yu, X. *et al.* Cryo-EM structure of human adenovirus D26 reveals the conservation of structural organization among human adenoviruses. *Sci. Adv.* **3**, e1602670 (2017).
47. Rux, J. J., Kuser, P. R. & Burnett, R. M. Structural and Phylogenetic Analysis of Adenovirus Hexons by Use of High-Resolution X-Ray Crystallographic, Molecular Modeling, and Sequence-Based Methods. *J. Virol.* **77**, 9553–9566 (2003).
48. Zubieta, C., Schoehn, G., Chroboczek, J. & Cusack, S. The Structure of the Human Adenovirus 2 Penton. *Mol. Cell* **17**, 121–135 (2005).
49. Wu, E. *et al.* Flexibility of the Adenovirus Fiber Is Required for Efficient Receptor Interaction. *J. Virol.* **77**, 7225–7235 (2003).
50. Gaggar, A., Shayakhmetov, D. M., Liszewski, M. K., Atkinson, J. P. & Lieber, A. Localization of Regions in CD46 That Interact with Adenovirus. *J. Virol.* **79**, 7503–7513 (2005).
51. Cupelli, K. *et al.* Structure of Adenovirus Type 21 Knob in Complex with CD46 Reveals Key Differences in Receptor Contacts among Species B Adenoviruses. *J. Virol.* **84**, 3189–3200 (2010).

52. Bewley, M. C., Springer, K., Zhang, Y.-B., Freimuth, P. & Flanagan, J. M. Structural Analysis of the Mechanism of Adenovirus Binding to Its Human Cellular Receptor, CAR. *Science* **286**, 1579–1583 (1999).
53. Xia, D., Henry, L. J., Gerard, R. D. & Deisenhofer, J. Crystal structure of the receptor-binding domain of adenovirus type 5 fiberprotein at 1.7 Å resolution. *Structure* **2**, 1259–1270 (1994).
54. Vassal-Stermann, E. *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).
55. Wang, H. *et al.* Desmoglein 2 is a receptor for adenovirus serotypes 3, 7, 11 and 14. *Nat. Med.* **17**, 96–104 (2011).
56. Arnberg, N., Edlund, K., Kidd, A. H. & Wadell, G. Adenovirus Type 37 Uses Sialic Acid as a Cellular Receptor. *J. Virol.* **74**, 42–48 (2000).
57. Lenman, A. *et al.* Polysialic acid is a cellular receptor for human adenovirus 52. *Proc. Natl. Acad. Sci.* 201716900 (2018) doi:10.1073/pnas.1716900115.
58. Baker, A. T., Mundy, R., Davies, J., Rizkallah, P. J. & Parker, A. L. Adenovirus serotype 26 utilises sialic acid bearing glycans as a primary cell entry receptor. *bioRxiv* 580076 (2019) doi:10.1101/580076.
59. Chandra, N. *et al.* Sialic Acid-Containing Glycans as Cellular Receptors for Ocular Human Adenoviruses: Implications for Tropism and Treatment. *Viruses* **11**, 395 (2019).
60. Nicol, C. G. *et al.* Effect of adenovirus serotype 5 fiber and penton modifications on in vivo tropism in rats. *Mol. Ther.* **10**, 344–354 (2004).

61. Bai, M., Harfe, B. & Freimuth, P. Mutations that alter an Arg-Gly-Asp (RGD) sequence in the adenovirus type 2 penton base protein abolish its cell-rounding activity and delay virus reproduction in flat cells. *J. Virol.* **67**, 5198–5205 (1993).
62. Parker, A. L. *et al.* Multiple vitamin K-dependent coagulation zymogens promote adenovirus-mediated gene delivery to hepatocytes. *Blood* **108**, 2554–2561 (2006).
63. Alba, R. *et al.* Identification of coagulation factor (F)X binding sites on the adenovirus serotype 5 hexon: effect of mutagenesis on FX interactions and gene transfer. *Blood* **114**, 965–971 (2009).
64. Russell, W. C. Adenoviruses: update on structure and function. *J. Gen. Virol.* **90**, 1–20 (2009).
65. Russell, S. J., Peng, K. W. & Bell, J. C. Oncolytic virotherapy. *Nat Biotechnol* **30**, 658–70 (2012).
66. Ghebremedhin, B. Human adenovirus: Viral pathogen with increasing importance. *Eur. J. Microbiol. Immunol.* **4**, 26–33 (2014).
67. Cohen, C. J. *et al.* The coxsackievirus and adenovirus receptor is a transmembrane component of the tight junction. *Proc. Natl. Acad. Sci.* **98**, 15191–15196 (2001).
68. Awasthi, V., Meinken, G., Springer, K., Srivastava, S. C. & Freimuth, P. Biodistribution of Radioiodinated Adenovirus Fiber Protein Knob Domain after Intravenous Injection in Mice. *J. Virol.* **78**, 6431–6438 (2004).
69. Bergelson, J. M. *et al.* Isolation of a Common Receptor for Coxsackie B Viruses and Adenoviruses 2 and 5. *Science* **275**, 1320–1323 (1997).

70. Roelvink, P. W. *et al.* 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 (1998).
71. Lenman, A. *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).
72. Baker, A. T. *et al.* Diversity within the adenovirus fiber knob hypervariable loops influences primary receptor interactions. *Nat. Commun.* **10**, 741 (2019).
73. Seiradake, E., Lortat-Jacob, H., Billet, O., Kremer, E. J. & Cusack, S. 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).
74. Roelvink, P. W., Mi Lee, G., Einfeld, D. A., Kovessi, I. & Wickham, T. J. Identification of a conserved receptor-binding site on the fiber proteins of CAR-recognizing adenoviridae. *Science* **286**, 1568–1571 (1999).
75. Kirby, I. *et al.* Identification of Contact Residues and Definition of the CAR-Binding Site of Adenovirus Type 5 Fiber Protein. *J. Virol.* **74**, 2804–2813 (2000).
76. Jakubczak, J. L. *et al.* Adenovirus Type 5 Viral Particles Pseudotyped with Mutagenized Fiber Proteins Show Diminished Infectivity of Coxsackie B-Adenovirus Receptor-Bearing Cells. *J. Virol.* **75**, 2972–2981 (2001).

77. Kirby, I. *et al.* 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 (1999).
78. Smith, T. *et al.* In Vivo Hepatic Adenoviral Gene Delivery Occurs Independently of the Coxsackievirus–Adenovirus Receptor. *Mol. Ther.* **5**, 770–779 (2002).
79. Uusi-Kerttula, H. *et al.* Pseudotyped $\alpha\beta 6$ integrin-targeted adenovirus vectors for ovarian cancer therapies. *Oncotarget* **7**, 27926–27937 (2016).
80. Man, Y. K. S. *et al.* The Novel Oncolytic Adenoviral Mutant Ad5-3 Δ -A20T Retargeted to $\alpha\beta 6$ Integrins Efficiently Eliminates Pancreatic Cancer Cells. *Mol. Cancer Ther.* **17**, 575–587 (2018).
81. Alemany, R. & Curiel, D. T. CAR-binding ablation does not change biodistribution and toxicity of adenoviral vectors. *Gene Ther.* **8**, 1347–1353 (2001).
82. Seiradake, E. *et al.* The cell adhesion molecule ‘CAR’ and sialic acid on human erythrocytes influence adenovirus in vivo biodistribution. *PLoS Pathog.* **5**, e1000277 (2009).
83. Carlisle, R. C. *et al.* Human erythrocytes bind and inactivate type 5 adenovirus by presenting Coxsackie virus-adenovirus receptor and complement receptor 1. *Blood* **113**, 1909–1918 (2009).
84. Andrews, P. W. *et al.* A human cell-surface antigen defined by a monoclonal antibody and controlled by a gene on human chromosome 1. *Ann. Hum. Genet.* **49**, 31–39 (1985).

85. Manchester, M., Liszewski, M. K., Atkinson, J. P. & Oldstone, M. B. Multiple isoforms of CD46 (membrane cofactor protein) serve as receptors for measles virus. *Proc. Natl. Acad. Sci.* **91**, 2161–2165 (1994).
86. Cattaneo, R. Four Viruses, Two Bacteria, and One Receptor: Membrane Cofactor Protein (CD46) as Pathogens' Magnet. *J. Virol.* **78**, 4385–4388 (2004).
87. Yamamoto, H., Fara, A. F., Dasgupta, P. & Kemper, C. CD46: The 'multitasker' of complement proteins. *Int. J. Biochem. Cell Biol.* **45**, 2808–2820 (2013).
88. Persson, B. D. *et al.* Structure of the Extracellular Portion of CD46 Provides Insights into Its Interactions with Complement Proteins and Pathogens. *PLOS Pathog.* **6**, e1001122 (2010).
89. Fleischli, C. *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).
90. Persson, B. D. *et al.* Adenovirus type 11 binding alters the conformation of its receptor CD46. *Nat. Struct. Mol. Biol.* **14**, 164–166 (2007).
91. Wang, H. *et al.* Identification of CD46 binding sites within the adenovirus serotype 35 fiber knob. *J. Virol.* **81**, 12785–12792 (2007).
92. Adams, W. C. *et al.* Adenovirus type-35 vectors block human CD4+ T-cell activation via CD46 ligation. *Proc. Natl. Acad. Sci.* **108**, 7499–7504 (2011).
93. Hay, J., Carter, D., Lieber, A. & Astier, A. L. Recombinant Ad35 adenoviral proteins as potent modulators of human T cell activation. *Immunology* (2014) doi:10.1111/imm.12391.

94. Iacobelli-Martinez, M., Nepomuceno, R. R., Connolly, J. & Nemerow, G. R. CD46-Utilizing Adenoviruses Inhibit C/EBP β -Dependent Expression of Proinflammatory Cytokines. *J. Virol.* **79**, 11259–11268 (2005).
95. Jiang, H. *et al.* Oncolytic adenovirus research evolution: from cell-cycle checkpoints to immune checkpoints. *Curr Opin Virol* **13**, 33–9 (2015).
96. Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. *Science* **357**, eaan2507 (2017).
97. Kuhn, I. *et al.* Directed Evolution Generates a Novel Oncolytic Virus for the Treatment of Colon Cancer. *PLOS ONE* **3**, e2409 (2008).
98. Branton, P. E. & Roopchand, D. E. The role of adenovirus E4orf4 protein in viral replication and cell killing. *Oncogene* **20**, 7855–7865 (2001).
99. Di, Y., Seymour, L. & Fisher, K. Activity of a group B oncolytic adenovirus (ColoAd1) in whole human blood. *Gene Ther.* **21**, 440–443 (2014).
100. Illingworth, S. *et al.* Preclinical Safety Studies of Enadenotucirev, a Chimeric Group B Human-Specific Oncolytic Adenovirus. *Mol. Ther. Oncolytics* **5**, 62–74 (2017).
101. Segerman, A. *et al.* Adenovirus Type 11 Uses CD46 as a Cellular Receptor. *J. Virol.* **77**, 9183–9191 (2003).
102. Human adenovirus B strain ColoAd1, complete genome. (2009).
103. Li, H. *et al.* Adenovirus Serotype 26 Utilizes CD46 as a Primary Cellular Receptor and Only Transiently Activates T Lymphocytes following Vaccination of Rhesus Monkeys. *J. Virol.* **86**, 10862–10865 (2012).

104. Baden, L. R. *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).
105. 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.
<https://clinicaltrials.gov/ct2/show/NCT02661464?term=adenovirus&recrs=a&cond=ebola&rank=1>.
106. Chen, H. *et al.* Adenovirus-Based Vaccines: Comparison of Vectors from Three Species of Adenoviridae. *J. Virol.* **84**, 10522–10532 (2010).
107. Uusi-Kerttula, H., Hulin-Curtis, S., Davies, J. & Parker, A. L. Oncolytic Adenovirus: Strategies and Insights for Vector Design and Immuno-Oncolytic Applications. *Viruses* **7**, 6009–6042 (2015).
108. Fender, P., Boussaid, A., Mezin, P. & Chroboczek, J. Synthesis, cellular localization, and quantification of penton-dodecahedron in serotype 3 adenovirus-infected cells. *Virology* **340**, 167–173 (2005).
109. Fuschiotti, P. *et al.* Structure of the dodecahedral penton particle from human adenovirus type 3. *J. Mol. Biol.* **356**, 510–520 (2006).
110. Wang, H. *et al.* Multimerization of Adenovirus Serotype 3 Fiber Knob Domains Is Required for Efficient Binding of Virus to Desmoglein 2 and Subsequent Opening of Epithelial Junctions. *J. Virol.* **85**, 6390–6402 (2011).

111. Lu, Z.-Z. *et al.* Penton-Dodecahedral Particles Trigger Opening of Intercellular Junctions and Facilitate Viral Spread during Adenovirus Serotype 3 Infection of Epithelial Cells. *PLOS Pathog.* **9**, e1003718 (2013).
112. Fender, P., Hall, K., Schoehn, G. & Blair, G. E. Impact of Human Adenovirus Type 3 Dodecahedron on Host Cells and Its Potential Role in Viral Infection. *J. Virol.* **86**, 5380–5385 (2012).
113. Murakami, M. *et al.* Chimeric Adenoviral Vectors Incorporating a Fiber of Human Adenovirus 3 Efficiently Mediate Gene Transfer into Prostate Cancer Cells. *The Prostate* **70**, 362–376 (2010).
114. Kim, K. H. *et al.* A phase I clinical trial of Ad5/3- Δ 24, a novel serotype-chimeric, infectivity-enhanced, conditionally-replicative adenovirus (CRAd), in patients with recurrent ovarian cancer. *Gynecol. Oncol.* **130**, 518–524 (2013).
115. Ranki, T. *et al.* Phase I study with ONCOS-102 for the treatment of solid tumors - an evaluation of clinical response and exploratory analyses of immune markers. *J. Immunother. Cancer* **4**, 17 (2016).
116. ONCOS-102 (Previously CGTG-102) for Therapy of Advanced Cancers - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT01598129>.
117. A Phase 1/2 Study to Investigate the Safety, Biologic and Anti-tumor Activity of ONCOS-102 in Combination With Durvalumab in Subjects With Advanced Peritoneal Malignancies - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT02963831>.

118. A Randomised Phase II Open-label Study With a Phase Ib Safety lead-in Cohort of ONCOS-102, an Immune-priming GM-CSF Coding Oncolytic Adenovirus, and Pemetrexed/Cisplatin in Patients With Unresectable Malignant Pleural Mesothelioma - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT02879669>.
119. Targovax announces completed enrollment of ONCOS-102 trial in mesothelioma. *Targovax* <https://www.targovax.com:443/en/targovax-announces-completed-enrollment-of-oncos-102-trial-in-mesothelioma/>.
120. Beyer, I. *et al.* Epithelial Junction Opener JO-1 Improves Monoclonal Antibody Therapy of Cancer. *Cancer Res.* **71**, 7080–7090 (2011).
121. Beyer, I. *et al.* Coadministration of Epithelial Junction Opener JO-1 Improves the Efficacy and Safety of Chemotherapeutic Drugs. *Clin. Cancer Res.* **18**, 3340–3351 (2012).
122. Richter, M. *et al.* Preclinical safety and efficacy studies with an affinity-enhanced epithelial junction opener and PEGylated liposomal doxorubicin. *Mol. Ther. - Methods Clin. Dev.* **2**, 15005 (2015).
123. Wang, H. *et al.* Structural and Functional Studies on the Interaction of Adenovirus Fiber Knobs and Desmoglein 2. *J. Virol.* **87**, 11346–11362 (2013).
124. Durmort, C. *et al.* Structure of the fiber head of Ad3, a non-CAR-binding serotype of adenovirus. *Virology* **285**, 302–312 (2001).
125. Vassal-Stermann, E. *et al.* CryoEM structure of adenovirus type 3 fibre with desmoglein 2 shows an unusual mode of receptor engagement. *Nat. Commun.* **10**, 1181 (2019).

126. Wang, H. *et al.* A New Human DSG2-Transgenic Mouse Model for Studying the Tropism and Pathology of Human Adenoviruses. *J. Virol.* **86**, 6286–6302 (2012).
127. Ye, X. *et al.* Upholding a role for EMT in breast cancer metastasis. *Nature* **547**, E1–E3 (2017).
128. Beerling, E. *et al.* Plasticity between Epithelial and Mesenchymal States Unlinks EMT from Metastasis-Enhancing Stem Cell Capacity. *Cell Rep.* **14**, 2281–2288 (2016).
129. Jolly, M. K., Ware, K. E., Gilja, S., Somarelli, J. A. & Levine, H. EMT and MET: necessary or permissive for metastasis? *Mol. Oncol.* **11**, 755–769 (2017).
130. Brabletz, T., Kalluri, R., Nieto, M. A. & Weinberg, R. A. EMT in cancer. *Nat. Rev. Cancer* **18**, 128–134 (2018).
131. Stencel-Baerenwald, J. E., Reiss, K., Reiter, D. M., Stehle, T. & Dermody, T. S. The sweet spot: defining virus–sialic acid interactions. *Nat. Rev. Microbiol.* **12**, 739–749 (2014).
132. Singh, A. K. *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).
133. Schnaar, R. L., Gerardy-Schahn, R. & Hildebrandt, H. Sialic Acids in the Brain: Gangliosides and Polysialic Acid in Nervous System Development, Stability, Disease, and Regeneration. *Physiol. Rev.* **94**, 461–518 (2014).

134. Colley, K. J., Kitajima, K. & Sato, C. Polysialic acid: Biosynthesis, novel functions and applications. *Crit. Rev. Biochem. Mol. Biol.* **49**, 498–532 (2014).
135. Rutishauser, U. Polysialic acid in the plasticity of the developing and adult vertebrate nervous system. *Nat. Rev. Neurosci.* **9**, 26–35 (2008).
136. Galuska, C. E., Lütteke, T. & Galuska, S. P. Is Polysialylated NCAM Not Only a Regulator during Brain Development But also during the Formation of Other Organs? *Biology* **6**, 27 (2017).
137. Kiermaier, E. *et al.* Polysialylation controls dendritic cell trafficking by regulating chemokine recognition. *Science* **351**, 186–190 (2016).
138. Jones, M. S. *et al.* New Adenovirus Species Found in a Patient Presenting with Gastroenteritis. *J. Virol.* **81**, 5978–5984 (2007).
139. Suzuki, M. *et al.* Polysialic acid facilitates tumor invasion by glioma cells. *Glycobiology* **15**, 887–894 (2005).
140. Petridis, A. K., Wedderkopp, H., Hugo, H. H. & Mehdorn, H. M. Polysialic acid overexpression in malignant astrocytomas. *Acta Neurochir. (Wien)* **151**, 601–604 (2009).
141. Amoureux, M.-C. *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).
142. Figarella-Branger, D. F., Durbec, P. L. & Rougon, G. N. Differential Spectrum of Expression of Neural Cell Adhesion Molecule Isoforms and L1 Adhesion Molecules on Human Neuroectodermal Tumors. *Cancer Res.* **50**, 6364–6370 (1990).

143. Glüer, S., Schelp, C., Gerardy-Schahn, R. & von Schweinitz, D. Polysialylated neural cell adhesion molecule as a marker for differential diagnosis in pediatric tumors. *J. Pediatr. Surg.* **33**, 1516–1520 (1998).
144. Lantuejoul, S., Moro, D., Michalides, R. J. a m, Brambilla, C. & Brambilla, E. Neural Cell Adhesion Molecules (ncam) and Ncam-*psa* Expression in Neuroendocrine Lung Tumors. *Am. J. Surg. Pathol.* **22**, 1267–1276 (1998).
145. Tanaka, F. *et al.* Expression of Polysialic Acid and STX, a Human Polysialyltransferase, Is Correlated with Tumor Progression in Non-Small Cell Lung Cancer. *Cancer Res.* **60**, 3072–3080 (2000).
146. Arnberg, N., Pring-Åkerblom, P. & Wadell, G. Adenovirus Type 37 Uses Sialic Acid as a Cellular Receptor on Chang C Cells. *J. Virol.* **76**, 8834–8841 (2002).
147. Nilsson, E. C. *et al.* The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. *Nat. Med.* **17**, 105–109 (2011).
148. Heneghan, C. J. *et al.* Zanamivir for influenza in adults and children: systematic review of clinical study reports and summary of regulatory comments. *BMJ* **348**, g2547 (2014).
149. McAuley, J. L., Gilbertson, B. P., Trifkovic, S., Brown, L. E. & McKimm-Breschkin, J. L. Influenza Virus Neuraminidase Structure and Functions. *Front. Microbiol.* **10**, (2019).
150. Schäfer, G., Blumenthal, M. J. & Katz, A. A. Interaction of Human Tumor Viruses with Host Cell Surface Receptors and Cell Entry. *Viruses* **7**, 2592–2617 (2015).

151. Bartlett, A. H. & Park, P. W. Proteoglycans in host–pathogen interactions: molecular mechanisms and therapeutic implications. *Expert Rev. Mol. Med.* **12**, e5 (2010).
152. Tuve, S. *et al.* Role of Cellular Heparan Sulfate Proteoglycans in Infection of Human Adenovirus Serotype 3 and 35. *PLOS Pathog.* **4**, e1000189 (2008).
153. Dechecchi, M. C., Tamanini, A., Bonizzato, A. & Cabrini, G. Heparan Sulfate Glycosaminoglycans Are Involved in Adenovirus Type 5 and 2-Host Cell Interactions. *Virology* **268**, 382–390 (2000).
154. Dechecchi, M. C. *et al.* Heparan Sulfate Glycosaminoglycans Are Receptors Sufficient To Mediate the Initial Binding of Adenovirus Types 2 and 5. *J. Virol.* **75**, 8772–8780 (2001).
155. Ruoslahti, E. Proteoglycans in cell regulation. *J. Biol. Chem.* **264**, 13369–13372 (1989).
156. Paolo, N. C. D., Kalyuzhniy, O. & Shayakhmetov, D. M. Fiber Shaft-Chimeric Adenovirus Vectors Lacking the KKTK Motif Efficiently Infect Liver Cells In Vivo. *J. Virol.* **81**, 12249–12259 (2007).
157. Kritz, A. B. *et al.* Adenovirus 5 Fibers Mutated at the Putative HSPG-binding Site Show Restricted Retargeting with Targeting Peptides in the HI Loop. *Mol. Ther.* **15**, 741–749 (2007).
158. Bayo-Puxan, N. *et al.* Replacement of Adenovirus Type 5 Fiber Shaft Heparan Sulfate Proteoglycan-Binding Domain with RGD for Improved Tumor Infectivity and Targeting. *Hum. Gene Ther.* **20**, 1214–1221 (2009).

159. Shayakhmetov, D. M., Gaggar, A., Ni, S., Li, Z.-Y. & Lieber, A. Adenovirus Binding to Blood Factors Results in Liver Cell Infection and Hepatotoxicity. *J. Virol.* **79**, 7478–7491 (2005).
160. Waddington, S. N. *et al.* Adenovirus Serotype 5 Hexon Mediates Liver Gene Transfer. *Cell* **132**, 397–409 (2008).
161. Kalyuzhniy, O. *et al.* Adenovirus serotype 5 hexon is critical for virus infection of hepatocytes in vivo. *Proc. Natl. Acad. Sci.* **105**, 5483–5488 (2008).
162. Vigant, F. *et al.* Substitution of Hexon Hypervariable Region 5 of Adenovirus Serotype 5 Abrogates Blood Factor Binding and Limits Gene Transfer to Liver. *Mol. Ther.* **16**, 1474–1480 (2008).
163. Alba, R. *et al.* Biodistribution and retargeting of FX-binding ablated adenovirus serotype 5 vectors. *Blood* **116**, 2656–2664 (2010).
164. Xu, Z. *et al.* Coagulation factor X shields adenovirus type 5 from attack by natural antibodies and complement. *Nat. Med.* **19**, 452–457 (2013).
165. Duffy, M. R., Doszpoly, A., Turner, G., Nicklin, S. A. & Baker, A. H. The relevance of coagulation factor X protection of adenoviruses in human sera. *Gene Ther.* **23**, 592–596 (2016).
166. Schmid, M. *et al.* Adenoviral vector with shield and adapter increases tumor specificity and escapes liver and immune control. *Nat. Commun.* **9**, 450 (2018).
167. Ma, J. *et al.* Manipulating Adenovirus Hexon Hypervariable Loops Dictates Immune Neutralisation and Coagulation Factor X-dependent Cell Interaction In Vitro and In Vivo. *PLoS Pathog.* **11**, (2015).

168. Alba, R. *et al.* Coagulation factor X mediates adenovirus type 5 liver gene transfer in non-human primates (*Microcebus murinus*). *Gene Ther.* **19**, 109–113 (2012).
169. Bradshaw, A. C. *et al.* Requirements for Receptor Engagement during Infection by Adenovirus Complexed with Blood Coagulation Factor X. *PLOS Pathog.* **6**, e1001142 (2010).
170. Corjon, S. *et al.* Cell Entry and Trafficking of Human Adenovirus Bound to Blood Factor X Is Determined by the Fiber Serotype and Not Hexon:Heparan Sulfate Interaction. *PLOS ONE* **6**, e18205 (2011).
171. Ranki, T. *et al.* A heparan sulfate-targeted conditionally replicative adenovirus, Ad5.pk7- Δ 24, for the treatment of advanced breast cancer. *Gene Ther.* **14**, 58–67 (2007).
172. Capasso, C. *et al.* Oncolytic adenoviruses coated with MHC-I tumor epitopes increase the antitumor immunity and efficacy against melanoma. *Oncoimmunology* **5**, (2015).
173. Konz, J. O. *et al.* Serotype Specificity of Adenovirus Purification Using Anion-Exchange Chromatography. *Hum. Gene Ther.* **16**, 1346–1353 (2005).
174. Wickham, T. J., Mathias, P., Cheresh, D. A. & Nemerow, G. R. Integrins $\alpha\beta 3$ and $\alpha\beta 5$ promote adenovirus internalization but not virus attachment. *Cell* **73**, 309–319 (1993).
175. Patterson, S. & Russell, W. C. Ultrastructural and Immunofluorescence Studies of Early Events in Adenovirus-HeLa Cell Interactions. *J. Gen. Virol.* **64**, 1091–1099 (1983).

176. Li, E., Stupack, D., Klemke, R., Cheresch, D. A. & Nemerow, G. R. Adenovirus Endocytosis via α v Integrins Requires Phosphoinositide-3-OH Kinase. *J. Virol.* **72**, 2055–2061 (1998).
177. Li, E., Stupack, D., Bokoch, G. M. & Nemerow, G. R. Adenovirus Endocytosis Requires Actin Cytoskeleton Reorganization Mediated by Rho Family GTPases. *J. Virol.* **72**, 8806–8812 (1998).
178. Bradshaw, A. C. *et al.* Biodistribution and inflammatory profiles of novel penton and hexon double-mutant serotype 5 adenoviruses. *J. Controlled Release* **164**, 394–402 (2012).
179. Henning, P. *et al.* Tumor cell targeted gene delivery by adenovirus 5 vectors carrying knobless fibers with antibody-binding domains. *Gene Ther.* **12**, 211–224 (2005).
180. Mathias, P., Wickham, T., Moore, M. & Nemerow, G. Multiple adenovirus serotypes use alpha v integrins for infection. *J. Virol.* **68**, 6811–6814 (1994).
181. Li, E. *et al.* Integrin α v β 1 Is an Adenovirus Coreceptor. *J. Virol.* **75**, 5405–5409 (2001).
182. Davison, E. *et al.* Adenovirus type 5 uptake by lung adenocarcinoma cells in culture correlates with Ad5 fibre binding is mediated by alpha(v)beta1 integrin and can be modulated by changes in beta1 integrin function. *J. Gene Med.* **3**, 550–559 (2001).
183. Huang, S., Kamata, T., Takada, Y., Ruggeri, Z. M. & Nemerow, G. R. Adenovirus interaction with distinct integrins mediates separate events in

- cell entry and gene delivery to hematopoietic cells. *J. Virol.* **70**, 4502–4508 (1996).
184. Davison, E., Diaz, R. M., Hart, I. R., Santis, G. & Marshall, J. F. Integrin alpha5beta1-mediated adenovirus infection is enhanced by the integrin-activating antibody TS2/16. *J. Virol.* **71**, 6204–6207 (1997).
185. Lindert, S., Silvestry, M., Mullen, T.-M., Nemerow, G. R. & Stewart, P. L. Cryo-Electron Microscopy Structure of an Adenovirus-Integrin Complex Indicates Conformational Changes in both Penton Base and Integrin. *J. Virol.* **83**, 11491–11501 (2009).
186. Meier, O. & Greber, U. F. Adenovirus endocytosis. *J. Gene Med.* **6**, S152–S163 (2004).
187. Mukherjee, S., Ghosh, R. N. & Maxfield, F. R. Endocytosis. *Physiol. Rev.* **77**, 759–803 (1997).
188. Wang, K., Guan, T., Cheresh, D. A. & Nemerow, G. R. Regulation of Adenovirus Membrane Penetration by the Cytoplasmic Tail of Integrin $\beta 5$. *J. Virol.* **74**, 2731–2739 (2000).
189. Chiu, C. Y., Mathias, P., Nemerow, G. R. & Stewart, P. L. Structure of Adenovirus Complexed with Its Internalization Receptor, $\alpha\beta 5$ Integrin. <http://jvi.asm.org>.
190. Acharya, R. *et al.* The three-dimensional structure of foot-and-mouth disease virus at 2.9 Å resolution. *Nature* **337**, 709–716 (1989).
191. Jackson, T. *et al.* Arginine-glycine-aspartic acid-specific binding by foot-and-mouth disease viruses to the purified integrin $\alpha(v)\beta 3$ in vitro. *J. Virol.* **71**, 8357–8361 (1997).

192. Lyle, C. & McCormick, F. Integrin $\alpha\beta 5$ is a primary receptor for adenovirus in CAR-negative cells. *Viol. J.* **7**, 148 (2010).
193. Cao, C. *et al.* Conserved fiber-penton base interaction revealed by nearly atomic resolution cryo-electron microscopy of the structure of adenovirus provides insight into receptor interaction. *J. Virol.* **86**, 12322–12329 (2012).
194. Nestić, D. *et al.* $\alpha\beta 3$ Integrin Is Required for Efficient Infection of Epithelial Cells with Human Adenovirus Type 26. *J. Virol.* **93**, e01474-18 (2019).
195. Lieber, A. *et al.* The role of Kupffer cell activation and viral gene expression in early liver toxicity after infusion of recombinant adenovirus vectors. *J. Virol.* **71**, 8798–8807 (1997).
196. Zaiss, A.-K. *et al.* Differential Activation of Innate Immune Responses by Adenovirus and Adeno-Associated Virus Vectors. *J. Virol.* **76**, 4580–4590 (2002).
197. Paolo, N. C. D. *et al.* IL-1 α and Complement Cooperate in Triggering Local Neutrophilic Inflammation in Response to Adenovirus and Eliminating Virus-Containing Cells. *PLOS Pathog.* **10**, e1004035 (2014).
198. Maler, M. D. *et al.* Key Role of the Scavenger Receptor MARCO in Mediating Adenovirus Infection and Subsequent Innate Responses of Macrophages. *mBio* **8**, e00670-17 (2017).
199. Stichling, N. *et al.* Lung macrophage scavenger receptor SR-A6 (MARCO) is an adenovirus type-specific virus entry receptor. *PLOS Pathog.* **14**, e1006914 (2018).
200. Lam, E., Stein, S. & Falck-Pedersen, E. Adenovirus Detection by the cGAS/STING/TBK1 DNA Sensing Cascade. *J. Virol.* **88**, 974–981 (2014).

201. Thorne, S. H. Adding STING to the Tale of Oncolytic Virotherapy. *Trends Cancer* **2**, 67–68 (2016).
202. Queiroz, N. M. G. P. de, Xia, T. & Barber, G. N. Defective STING signaling in ovarian cancer cells favor oncolytic virus action. *J. Immunol.* **198**, 130.28-130.28 (2017).
203. Konno, H. *et al.* Suppression of STING signaling through epigenetic silencing and missense mutation impedes DNA damage mediated cytokine production. *Oncogene* **37**, 2037–2051 (2018).
204. Myhre, S. *et al.* Decreased immune reactivity towards a knobless, affibody-targeted adenovirus type 5 vector. *Gene Ther.* **14**, 376–381 (2007).
205. Bradley, R. R. *et al.* Adenovirus Serotype 5-Specific Neutralizing Antibodies Target Multiple Hexon Hypervariable Regions. *J. Virol.* **86**, 1267–1272 (2012).
206. Bradley, R. R., Lynch, D. M., Iampietro, M. J., Borducchi, E. N. & Barouch, D. H. Adenovirus Serotype 5 Neutralizing Antibodies Target both Hexon and Fiber following Vaccination and Natural Infection. *J. Virol.* **86**, 625–629 (2012).
207. Wirtz, P. & Steipe, B. Intrabody construction and expression III: engineering hyperstable V(H) domains. *Protein Sci. Publ. Protein Soc.* **8**, 2245–2250 (1999).
208. Magnusson, M. K., Hong, S. S., Henning, P., Boulanger, P. & Lindholm, L. Genetic retargeting of adenovirus vectors: functionality of targeting ligands and their influence on virus viability. *J. Gene Med.* **4**, 356–370 (2002).

209. Kim, J. W. *et al.* A novel single-chain antibody redirects adenovirus to IL13R α 2-expressing brain tumors. *Sci. Rep.* **5**, 18133 (2015).
210. Poulin, K. L. *et al.* Retargeting of Adenovirus Vectors through Genetic Fusion of a Single-Chain or Single-Domain Antibody to Capsid Protein IX. *J. Virol.* **84**, 10074–10086 (2010).
211. Glasgow, J. N., Mikheeva, G., Krasnykh, V. & Curiel, D. T. A Strategy for Adenovirus Vector Targeting with a Secreted Single Chain Antibody. *PLOS ONE* **4**, e8355 (2009).
212. Nettelbeck, D. M. *et al.* Targeting of Adenovirus to Endothelial Cells by a Bispecific Single-Chain Diabody Directed against the Adenovirus Fiber Knob Domain and Human Endoglin (CD105). *Mol. Ther.* **3**, 882–891 (2001).
213. Kashentseva, E. A., Seki, T., Curiel, D. T. & Dmitriev, I. P. Adenovirus targeting to c-erbB-2 oncoprotein by single-chain antibody fused to trimeric form of adenovirus receptor ectodomain. *Cancer Res.* **62**, 609–616 (2002).
214. Ståhl, S. *et al.* Affibody Molecules in Biotechnological and Medical Applications. *Trends Biotechnol.* **35**, 691–712 (2017).
215. Frejd, F. Y. & Kim, K.-T. Affibody molecules as engineered protein drugs. *Exp. Mol. Med.* **49**, e306 (2017).
216. Henning, P. *et al.* Genetic modification of adenovirus 5 tropism by a novel class of ligands based on a three-helix bundle scaffold derived from staphylococcal protein A. *Hum. Gene Ther.* **13**, 1427–1439 (2002).

217. Magnusson, M. K. *et al.* Adenovirus 5 vector genetically re-targeted by an Affibody molecule with specificity for tumor antigen HER2/neu. *Cancer Gene Ther.* **14**, 468–479 (2007).
218. Belousova, N., Mikheeva, G., Gelovani, J. & Krasnykh, V. Modification of Adenovirus Capsid with a Designed Protein Ligand Yields a Gene Vector Targeted to a Major Molecular Marker of Cancer. *J. Virol.* **82**, 630–637 (2008).
219. Magnusson, M. K. *et al.* A Transductionally Retargeted Adenoviral Vector for Virotherapy of Her2/neu-Expressing Prostate Cancer. *Hum. Gene Ther.* **23**, 70–82 (2012).
220. Bourgeois-Daigneault, M.-C. *et al.* Neoadjuvant oncolytic virotherapy before surgery sensitizes triple-negative breast cancer to immune checkpoint therapy. *Sci. Transl. Med.* **10**, eaao1641 (2018).
221. Samson, A. *et al.* Intravenous delivery of oncolytic reovirus to brain tumor patients immunologically primes for subsequent checkpoint blockade. *Sci. Transl. Med.* **10**, eaam7577 (2018).
222. Lee, P. & Gujar, S. Potentiating prostate cancer immunotherapy with oncolytic viruses. *Nat. Rev. Urol.* **15**, 235–250 (2018).
223. Hamid, O., Hoffner, B., Gasal, E., Hong, J. & Carvajal, R. D. Oncolytic immunotherapy: unlocking the potential of viruses to help target cancer. *Cancer Immunol. Immunother.* **66**, 1249–1264 (2017).
224. Myhre, S. *et al.* Re-targeted adenovirus vectors with dual specificity; binding specificities conferred by two different Affibody molecules in the fiber. *Gene Ther.* **16**, 252–261 (2009).

225. Belousova, N. *et al.* Native and engineered tropism of vectors derived from a rare species D adenovirus serotype 43. *Oncotarget* **7**, 53414–53429 (2016).
226. Vrij, J. de *et al.* A cathepsin-cleavage site between the adenovirus capsid protein IX and a tumor-targeting ligand improves targeted transduction. *Gene Ther.* **19**, 899–906 (2012).
227. Vellinga, J., Wollenberg, D. J. M. van den, Heijdt, S. van der, Rabelink, M. J. W. E. & Hoeben, R. C. The Coiled-Coil Domain of the Adenovirus Type 5 Protein IX Is Dispensable for Capsid Incorporation and Thermostability. *J. Virol.* **79**, 3206–3210 (2005).
228. Meulenbroek, R. A., Sargent, K. L., Lunde, J., Jasmin, B. J. & Parks, R. J. Use of adenovirus protein IX (pIX) to display large polypeptides on the virion—generation of fluorescent virus through the incorporation of pIX-GFP. *Mol. Ther.* **9**, 617–624 (2004).
229. Campos, S. K., Parrott, M. B. & Barry, M. A. Avidin-based targeting and purification of a protein IX-modified, metabolically biotinylated adenoviral vector. *Mol. Ther.* **9**, 942–954 (2004).
230. Vellinga, J. *et al.* Spacers Increase the Accessibility of Peptide Ligands Linked to the Carboxyl Terminus of Adenovirus Minor Capsid Protein IX. *J. Virol.* **78**, 3470–3479 (2004).
231. Greber, U. F., Willetts, M., Webster, P. & Helenius, A. Stepwise dismantling of adenovirus 2 during entry into cells. *Cell* **75**, 477–486 (1993).

232. Campos, S. K. & Barry, M. A. Comparison of adenovirus fiber, protein IX, and hexon capsomeres as scaffolds for vector purification and cell targeting. *Virology* **349**, 453–462 (2006).
233. Turk, V. *et al.* Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1824**, 68–88 (2012).
234. Binz, H. K. *et al.* High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat. Biotechnol.* **22**, 575–582 (2004).
235. Steiner, D., Forrer, P. & Plückthun, A. Efficient Selection of DARPins with Sub-nanomolar Affinities using SRP Phage Display. *J. Mol. Biol.* **382**, 1211–1227 (2008).
236. Plückthun, A. Designed Ankyrin Repeat Proteins (DARPins): Binding Proteins for Research, Diagnostics, and Therapy. *Annu. Rev. Pharmacol. Toxicol.* **55**, 489–511 (2015).
237. Guillard, S. *et al.* Structural and functional characterization of a DARPin which inhibits Ras nucleotide exchange. *Nat. Commun.* **8**, 16111 (2017).
238. Dreier, B. *et al.* Her2-specific Multivalent Adapters Confer Designed Tropism to Adenovirus for Gene Targeting. *J. Mol. Biol.* **405**, 410–426 (2011).
239. Dreier, B. *et al.* Development of a generic adenovirus delivery system based on structure-guided design of bispecific trimeric DARPin adapters. *Proc. Natl. Acad. Sci.* **110**, E869–E877 (2013).

240. Sebestyén, Z., Vrij, J. de, Magnusson, M., Debets, R. & Willemsen, R. An Oncolytic Adenovirus Redirected with a Tumor-Specific T-Cell Receptor. *Cancer Res.* **67**, 11309–11316 (2007).
241. Meek, D. W. & Marcar, L. MAGE-A antigens as targets in tumour therapy. *Cancer Lett.* **324**, 126–132 (2012).
242. Zajac, P. *et al.* MAGE-A Antigens and Cancer Immunotherapy. *Front. Med.* **4**, (2017).
243. Ping, Y., Liu, C. & Zhang, Y. T-cell receptor-engineered T cells for cancer treatment: current status and future directions. *Protein Cell* **9**, 254–266 (2018).
244. Belousova, N., Krendelchtchikova, V., Curiel, D. T. & Krasnykh, V. Modulation of Adenovirus Vector Tropism via Incorporation of Polypeptide Ligands into the Fiber Protein. *J. Virol.* **76**, 8621–8631 (2002).
245. Uusi-Kerttula, H. *et al.* Incorporation of Peptides Targeting EGFR and FGFR1 into the Adenoviral Fiber Knob Domain and Their Evaluation as Targeted Cancer Therapies. *Hum. Gene Ther.* **26**, 320–329 (2015).
246. Magnusson, M. K., Hong, S. S., Boulanger, P. & Lindholm, L. Genetic Retargeting of Adenovirus: Novel Strategy Employing “Deknobbing” of the Fiber. *J. Virol.* **75**, 7280–7289 (2001).
247. Miura, Y. *et al.* Direct selection of targeted adenovirus vectors by random peptide display on the fiber knob. *Gene Ther.* **14**, 1448–1460 (2007).
248. Nishimoto, T. *et al.* Oncolytic virus therapy for pancreatic cancer using the adenovirus library displaying random peptides on the fiber knob. *Gene Ther.* **16**, 669–680 (2009).

249. Böckmann, M., Drosten, M. & Pützer, B. M. Discovery of targeting peptides for selective therapy of medullary thyroid carcinoma. *J. Gene Med.* **7**, 179–188 (2005).
250. Xia, H., Anderson, B., Mao, Q. & Davidson, B. L. Recombinant Human Adenovirus: Targeting to the Human Transferrin Receptor Improves Gene Transfer to Brain Microcapillary Endothelium. *J. Virol.* **74**, 11359–11366 (2000).
251. Beatty, M. S. & Curiel, D. T. Adenovirus Strategies for Tissue-Specific Targeting. *Adv. Cancer Res.* **115**, 39–67 (2012).
252. Bandyopadhyay, A. & Raghavan, S. Defining the Role of Integrin $\alpha\beta 6$ in Cancer. *Curr. Drug Targets* **10**, 645–652 (2009).
253. Patsenker, E. *et al.* The $\alpha\beta 6$ integrin is a highly specific immunohistochemical marker for cholangiocarcinoma. *J. Hepatol.* **52**, 362–369 (2010).
254. Kotecha, A. *et al.* Rules of engagement between $\alpha\beta 6$ integrin and foot-and-mouth disease virus. *Nat. Commun.* **8**, 15408 (2017).
255. Logan, D. *et al.* Structure of a major immunogenic site on foot-and-mouth disease virus. *Nature* **362**, 566–568 (1993).
256. DiCara, D. *et al.* Structure-Function Analysis of Arg-Gly-Asp Helix Motifs in $\alpha\beta 6$ Integrin Ligands. *J. Biol. Chem.* **282**, 9657–9665 (2007).
257. Hausner, S. H., DiCara, D., Marik, J., Marshall, J. F. & Sutcliffe, J. L. Use of a Peptide Derived from Foot-and-Mouth Disease Virus for the Noninvasive Imaging of Human Cancer: Generation and Evaluation of 4-[¹⁸F]Fluorobenzoyl A20FMDV2 for In vivo Imaging of Integrin $\alpha\beta 6$

- Expression with Positron Emission Tomography. *Cancer Res.* **67**, 7833–7840 (2007).
258. Coughlan, L. *et al.* In Vivo Retargeting of Adenovirus Type 5 to $\alpha\beta 6$ Integrin Results in Reduced Hepatotoxicity and Improved Tumor Uptake following Systemic Delivery. *J. Virol.* **83**, 6416–6428 (2009).
259. Coughlan, L. *et al.* Combined Fiber Modifications Both to Target $\alpha\beta 6$ and Detarget the Coxsackievirus–Adenovirus Receptor Improve Virus Toxicity Profiles In Vivo but Fail to Improve Antitumoral Efficacy Relative to Adenovirus Serotype 5. *Hum. Gene Ther.* **23**, 960–979 (2012).
260. Coughlan, L. *et al.* Retargeting Adenovirus Serotype 48 Fiber Knob Domain by Peptide Incorporation. *Hum. Gene Ther.* **25**, 385–394 (2014).
261. Öberg, D. *et al.* Improved potency and selectivity of an oncolytic E1ACR2 and E1B19K deleted adenoviral mutant (Ad $\Delta\Delta$) in prostate and pancreatic cancers. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **16**, 541–553 (2010).
262. Uusi-Kerttula, H. *et al.* Ad5NULL-A20 - a tropism-modified, $\alpha\beta 6$ integrin-selective oncolytic adenovirus for epithelial ovarian cancer therapies. *Clin. Cancer Res.* clincanres.1089.2018 (2018) doi:10.1158/1078-0432.CCR-18-1089.
263. Sela-Culang, I., Kunik, V. & Ofran, Y. The Structural Basis of Antibody-Antigen Recognition. *Front. Immunol.* **4**, (2013).
264. Stumpp, M. T. & Amstutz, P. DARPins: a true alternative to antibodies. *Curr. Opin. Drug Discov. Devel.* **10**, 153–159 (2007).

265. Huang, P.-S., Boyken, S. E. & Baker, D. The coming of age of *de novo* protein design. *Nature* **537**, 320–327 (2016).
266. Heise, C. *et al.* ONYX-015, an E1B gene-attenuated adenovirus, causes tumor-specific cytolysis and antitumoral efficacy that can be augmented by standard chemotherapeutic agents. *Nat Med* **3**, 639–45 (1997).
267. Ries, S. J. *et al.* Loss of p14ARF in tumor cells facilitates replication of the adenovirus mutant dl1520 (ONYX-015). *Nat Med* **6**, 1128–33 (2000).
268. Rivlin, N., Brosh, R., Oren, M. & Rotter, V. Mutations in the p53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis. *Genes Cancer* **2**, 466–74 (2011).
269. Xu, R.-H. *et al.* [Phase II clinical study of intratumoral H101, an E1B deleted adenovirus, in combination with chemotherapy in patients with cancer]. *Ai Zheng Aizheng Chin. J. Cancer* **22**, 1307–1310 (2003).
270. Lu, W. *et al.* Intra-tumor injection of H101, a recombinant adenovirus, in combination with chemotherapy in patients with advanced cancers: a pilot phase II clinical trial. *World J. Gastroenterol.* **10**, 3634–3638 (2004).
271. Xia, Z. J. *et al.* [Phase III randomized clinical trial of intratumoral injection of E1B gene-deleted adenovirus (H101) combined with cisplatin-based chemotherapy in treating squamous cell cancer of head and neck or esophagus.]. *Ai Zheng* **23**, 1666–70 (2004).
272. White, E. Mechanisms of apoptosis regulation by viral oncogenes in infection and tumorigenesis. *Cell Death Differ* **13**, 1371–7 (2006).
273. Ahmed, A. A. *et al.* Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J. Pathol.* **221**, 49–56 (2010).

274. Olivier, M., Hollstein, M. & Hainaut, P. TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harb. Perspect. Biol.* **2**, (2010).
275. Goodrum, F. D. & Ornelles, D. A. p53 status does not determine outcome of E1B 55-kilodalton mutant adenovirus lytic infection. *J Virol* **72**, 9479–90 (1998).
276. Elsing, A. & Burgert, H.-G. The adenovirus E3/10.4K–14.5K proteins down-modulate the apoptosis receptor Fas/Apo-1 by inducing its internalization. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10072–10077 (1998).
277. Cheong, S. C. *et al.* E1A-expressing adenoviral E3B mutants act synergistically with chemotherapeutics in immunocompetent tumor models. *Cancer Gene Ther* **15**, 40–50 (2008).
278. O’Shea, C. C. *et al.* Late viral RNA export, rather than p53 inactivation, determines ONYX-015 tumor selectivity. *Cancer Cell* **6**, 611–23 (2004).
279. Liang, M. Oncorine, the World First Oncolytic Virus Medicine and its Update in China. *Curr Cancer Drug Targets* **18**, 171–176 (2018).
280. Berk, A. J. Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. *Oncogene* **24**, 7673–85 (2005).
281. Heise, C. *et al.* An adenovirus E1A mutant that demonstrates potent and selective systemic anti-tumoral efficacy. *Nat Med* **6**, 1134–9 (2000).
282. Fueyo, J. *et al.* A mutant oncolytic adenovirus targeting the Rb pathway produces anti-glioma effect in vivo. *Oncogene* **19**, 2–12 (2000).

283. Lang, F. F. *et al.* Phase I Study of DNX-2401 (Delta-24-RGD) Oncolytic Adenovirus: Replication and Immunotherapeutic Effects in Recurrent Malignant Glioma. *J Clin Oncol* **36**, 1419–1427 (2018).
284. Kuryk, L. *et al.* Toxicological and bio-distribution profile of a GM-CSF-expressing, double-targeted, chimeric oncolytic adenovirus ONCOS-102 - Support for clinical studies on advanced cancer treatment. *PLoS One* **12**, e0182715 (2017).
285. Rojas, J. J. *et al.* A modified E2F-1 promoter improves the efficacy to toxicity ratio of oncolytic adenoviruses. *Gene Ther* **16**, 1441–51 (2009).
286. Rodriguez-Garcia, A. *et al.* Safety and efficacy of VCN-01, an oncolytic adenovirus combining fiber HSG-binding domain replacement with RGD and hyaluronidase expression. *Clin Cancer Res* **21**, 1406–18 (2015).
287. A Phase I Dose Escalation Study of Intratumoral VCN-01 Injections With Gemcitabine and Abraxane® in Patients With Advanced Pancreatic Cancer - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT02045589>.
288. Phase I Dose Escalation Study of Intravenous VCN-01 With or Without Gemcitabine and Abraxane® in Patients With Advanced Solid Tumors - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT02045602>.
289. Burgert, H. G. & Kvist, S. An adenovirus type 2 glycoprotein blocks cell surface expression of human histocompatibility class I antigens. *Cell* **41**, 987–997 (1985).

290. Dong, W. *et al.* ORCA-010, a novel potency-enhanced oncolytic adenovirus, exerts strong antitumor activity in preclinical models. *Hum Gene Ther* **25**, 897–904 (2014).
291. Dyer, A. *et al.* Oncolytic Group B Adenovirus Enadenotucirev Mediates Non-apoptotic Cell Death with Membrane Disruption and Release of Inflammatory Mediators. *Mol Ther Oncolytics* **4**, 18–30 (2017).
292. Bressy, C. & Benihoud, K. Association of oncolytic adenoviruses with chemotherapies: an overview and future directions. *Biochem Pharmacol* **90**, 97–106 (2014).
293. Rodriguez, R. *et al.* Prostate Attenuated Replication Competent Adenovirus (ARCA) CN706: A Selective Cytotoxic for Prostate-specific Antigen-positive Prostate Cancer Cells. *Cancer Res.* **57**, 2559–2563 (1997).
294. Huang, T.-G., Savontaus, M. J., Shinozaki, K., Sauter, B. V. & Woo, S. L. C. Telomerase-dependent oncolytic adenovirus for cancer treatment. *Gene Ther.* **10**, 1241–1247 (2003).
295. Nemunaitis, J. *et al.* A phase I study of telomerase-specific replication competent oncolytic adenovirus (telomelysin) for various solid tumors. *Mol Ther* **18**, 429–34 (2010).
296. Dilley, J. *et al.* Oncolytic adenovirus CG7870 in combination with radiation demonstrates synergistic enhancements of antitumor efficacy without loss of specificity. *Cancer Gene Ther* **12**, 715–22 (2005).
297. Small, E. J. *et al.* A phase I trial of intravenous CG7870, a replication-selective, prostate-specific antigen-targeted oncolytic adenovirus, for the

treatment of hormone-refractory, metastatic prostate cancer. *Mol Ther* **14**, 107–17 (2006).

298. Efficacy & Safety of rAd-IFN Administered With Celecoxib & Gemcitabine in Patients With Malignant Pleural Mesothelioma - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT03710876>.

299. A Study to Evaluate INSTILADRIN® in Patients With High-Grade, Bacillus Calmette-Guerin (BCG) Unresponsive NMIBC - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT02773849>.

300. Multicenter RCT of ADV-TK Gene Therapy Improving the Outcome of Liver Transplantation for Advanced HCC - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT03313596>.

301. Tyynelä, K. *et al.* Adenovirus-mediated herpes simplex virus thymidine kinase gene therapy in BT4C rat glioma model. *Cancer Gene Ther.* **9**, 917–924 (2002).

302. Eastham, J. A. *et al.* Prostate cancer gene therapy: herpes simplex virus thymidine kinase gene transduction followed by ganciclovir in mouse and human prostate cancer models. *Hum Gene Ther* **7**, 515–23 (1996).

303. Nagabhushan, T. L. *et al.* Enhancement of intravesical delivery with Syn3 potentiates interferon-alpha2b gene therapy for superficial bladder cancer. *Cytokine Growth Factor Rev.* **18**, 389–394 (2007).

304. HAIC Plus H101 vs HAIC Alone for Unresectable HCC at BCLC A-B - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT03780049>.

305. Wennier, S. T., Liu, J. & McFadden, G. Bugs and Drugs: Oncolytic Virotherapy in Combination with Chemotherapy. *Curr. Pharm. Biotechnol.* **13**, 1817–1833 (2012).
306. LOAd703 Oncolytic Virus Therapy for Pancreatic Cancer - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT02705196>.
307. Trial Investigating an Immunostimulatory Oncolytic Adenovirus for Cancer - Full Text View - ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT03225989>.
308. Eriksson, E. *et al.* Shaping the Tumor Stroma and Sparking Immune Activation by CD40 and 4-1BB Signaling Induced by an Armed Oncolytic Virus. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **23**, 5846–5857 (2017).
309. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).
310. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539–539 (2011).
311. Li, W. *et al.* The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* **43**, W580-4 (2015).
312. Garcia-Boronat, M., Diez-Rivero, C. M., Reinherz, E. L. & Reche, P. A. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Res.* **36**, W35–W41 (2008).

313. Jukes, T. H. & Cantor, C. R. Evolution of Protein Molecules. in *Mammalian Protein Metabolism* 21–132 (Elsevier, 1969). doi:10.1016/B978-1-4832-3211-9.50009-7.
314. Zuckerkandl, E. & Pauling, L. Evolutionary Divergence and Convergence in Proteins. in *Evolving Genes and Proteins* (eds. Bryson, V. & Vogel, H. J.) 97–166 (Academic Press, 1965). doi:10.1016/B978-1-4832-2734-4.50017-6.
315. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
316. Felsenstein, J. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evol. Int. J. Org. Evol.* **39**, 783–791 (1985).
317. Stanton, R. J., McSharry, B. P., Armstrong, M., Tomasec, P. & Wilkinson, G. W. G. Re-engineering adenovirus vector systems to enable high-throughput analyses of gene function. *BioTechniques* **45**, 659–662, 664–668 (2008).
318. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
319. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).
320. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).

321. Kaneko, H. *et al.* Analysis of the complete genome sequence of epidemic keratoconjunctivitis-related human adenovirus type 8, 19, 37 and a novel serotype. *J. Gen. Virol.* **90**, 1471–1476 (2009).
322. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 2.0.
323. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
324. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
325. Winter, G. *et al.* DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 85–97 (2018).
326. Vonrhein, C. *et al.* Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 293–302 (2011).
327. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006).
328. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
329. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292 (2011).
330. Dodson, E. J., Winn, M. & Ralph, A. [32] Collaborative computational project, number 4: Providing programs for protein crystallography. in *Methods in Enzymology* vol. 277 620–633 (Academic Press, 1997).
331. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

332. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
333. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
334. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665-667 (2004).
335. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10037–10041 (2001).
336. Krieger, E. *et al.* Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins Struct. Funct. Bioinforma.* **77**, 114–122 (2009).
337. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
338. Jensen, J. H. Predicting accurate absolute binding energies in aqueous solution: thermodynamic considerations for electronic structure methods. *Phys. Chem. Chem. Phys.* **17**, 12441–12451 (2015).
339. Baker, A. T. *et al.* Diversity within the adenovirus fiber knob hypervariable loops influences primary receptor interactions. *bioRxiv* 406819 (2018) doi:10.1101/406819.
340. Pauly, M. *et al.* High prevalence and diversity of species D adenoviruses (HAdV-D) in human populations of four Sub-Saharan countries. *Viol. J.* **11**, 25 (2014).

341. Milligan, I. D. *et al.* Safety and Immunogenicity of Novel Adenovirus Type 26- and Modified Vaccinia Ankara-Vectored Ebola Vaccines: A Randomized Clinical Trial. *JAMA* **315**, 1610–1623 (2016).
342. Naso, M. F., Tomkowicz, B., Perry, W. L. & Strohl, W. R. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *Biodrugs* **31**, 317–334 (2017).
343. Schnurr, D. & Dondero, M. E. Two New Candidate Adenovirus Serotypes. *Intervirology* **36**, 79–83 (1993).
344. Sarantis, H., Johnson, G., Brown, M., Petric, M. & Tellier, R. Comprehensive Detection and Serotyping of Human Adenoviruses by PCR and Sequencing. *J. Clin. Microbiol.* **42**, 3963–3969 (2004).
345. Ishiko, H. *et al.* Novel Human Adenovirus Causing Nosocomial Epidemic Keratoconjunctivitis. *J. Clin. Microbiol.* **46**, 2002–2008 (2008).
346. Human adenovirus 49 gene for fiber protein, partial cds, strain: T87-677. (2016).
347. Al Qurashi, Y. M. A., Alkhalaf, M. A., Lim, L., Guiver, M. & Cooper, R. J. Sequencing and phylogenetic analysis of the hexon, fiber, and penton regions of adenoviruses isolated from AIDS patients. *J. Med. Virol.* **84**, 1157–1165 (2012).
348. De Jong, J. C. *et al.* Adenoviruses from Human Immunodeficiency Virus-Infected Individuals, Including Two Strains That Represent New Candidate Serotypes Ad50 and Ad51 of Species B1 and D, Respectively. *J. Clin. Microbiol.* **37**, 3940–3945 (1999).

349. Vogels, R. *et al.* Replication-Deficient Human Adenovirus Type 35 Vectors for Gene Transfer and Vaccination: Efficient Human Cell Infection and Bypass of Preexisting Adenovirus Immunity. *J. Virol.* **77**, 8263–8271 (2003).
350. Dakin, R. S., Parker, A. L., Delles, C., Nicklin, S. A. & Baker, A. H. Efficient transduction of primary vascular cells by the rare adenovirus serotype 49 vector. *Hum. Gene Ther.* **26**, 312–319 (2015).
351. Lemckert, A. A. C. *et al.* Generation of a novel replication-incompetent adenoviral vector derived from human adenovirus type 49: manufacture on PER.C6 cells, tropism and immunogenicity. *J. Gen. Virol.* **87**, 2891–2899 (2006).
352. Sharma, P., Martis, P. C. & Excoffon, K. J. D. A. Adenovirus transduction: More complicated than receptor expression. *Virology* **502**, 144–151 (2017).
353. Butterfield, L. H. & Vujanovic, L. New approaches to the development of adenoviral dendritic cell vaccines in melanoma. *Curr. Opin. Investig. Drugs Lond. Engl.* **2000 11**, 1399–1408 (2010).
354. Delva, E., Tucker, D. K. & Kowalczyk, A. P. The Desmosome. *Cold Spring Harb. Perspect. Biol.* **1**, (2009).
355. Robinson, C. M., Seto, D., Jones, M. S., Dyer, D. W. & Chodosh, J. Molecular evolution of human species D adenoviruses. *Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis.* **11**, 1208–1217 (2011).
356. Baker, A. T., Mundy, R. M., Davies, J. A., Rizkallah, P. J. & Parker, A. L. Human adenovirus type 26 uses sialic acid-bearing glycans as a primary cell entry receptor. *Sci. Adv.* **5**, eaax3567 (2019).

357. Tcatchoff, L. *et al.* Annexin A1 and A2: Roles in Retrograde Trafficking of Shiga Toxin. *PLOS ONE* **7**, e40429 (2012).
358. Lauritzen, S. P., Boye, T. L. & Nylandsted, J. Annexins are instrumental for efficient plasma membrane repair in cancer cells. *Semin. Cell Dev. Biol.* **45**, 32–38 (2015).
359. Mikuličić, S. & Florin, L. The endocytic trafficking pathway of oncogenic papillomaviruses. *Papillomavirus Res.* **7**, 135 (2019).
360. Bryksin, A. & Laktionov, P. Role of glyceraldehyde-3-phosphate dehydrogenase in vesicular transport from golgi apparatus to endoplasmic reticulum. - PubMed - NCBI. <https://www.ncbi.nlm.nih.gov/pubmed/18620527>.
361. Duncan, P. I., Stojdl, D. F., Marius, R. M., Scheit, K. H. & Bell, J. C. The Clk2 and Clk3 Dual-Specificity Protein Kinases Regulate the Intranuclear Distribution of SR Proteins and Influence Pre-mRNA Splicing. *Exp. Cell Res.* **241**, 300–308 (1998).
362. Ozawa, M., Terada, H. & Pedraza, C. The Fourth Armadillo Repeat of Plakoglobin (γ -Catenin) Is Required for Its High Affinity Binding to the Cytoplasmic Domains of E-Cadherin and Desmosomal Cadherin Dsg2, and the Tumor Suppressor APC Protein. *J. Biochem. (Tokyo)* **118**, 1077–1082 (1995).
363. JK3rd *et al.* The amino- and carboxyl-terminal tails of (beta)-catenin reduce its affinity for desmoglein 2. *J. Cell Sci.* **113**, 1737–1745 (2000).
364. Chandra, N. *et al.* Sulfated Glycosaminoglycans as Viral Decoy Receptors for Human Adenovirus Type 37. *Viruses* **11**, 247 (2019).

365. Chandra, N., Frångsmyr, L. & Arnberg, N. Decoy Receptor Interactions as Novel Drug Targets against EKC-Causing Human Adenovirus. *Viruses* **11**, 242 (2019).
366. Shin, K.-S. *et al.* Monocyte-Derived Dendritic Cells Dictate the Memory Differentiation of CD8+ T Cells During Acute Infection. *Front. Immunol.* **10**, 1887 (2019).
367. Wculek, S. K. *et al.* Dendritic cells in cancer immunology and immunotherapy. *Nat. Rev. Immunol.* (2019) doi:10.1038/s41577-019-0210-z.
368. Chen, M. & Wang, J. Programmed cell death of dendritic cells in immune regulation. *Immunol. Rev.* **236**, 11–27 (2010).
369. A Study to Assess Ebola Vaccines ChAd3-EBO-Z and Ad26.ZEBOV - Full Text View - ClinicalTrials.gov.
<https://clinicaltrials.gov/ct2/show/NCT02495246?term=ad26&draw=1&rank=2>.
370. Gaggar, A., Shayakhmetov, D. M. & Lieber, A. CD46 is a cellular receptor for group B adenoviruses. *Nat. Med.* **9**, 1408–1412 (2003).
371. Pache, L., Venkataraman, S., Reddy, V. S. & Nemerow, G. R. Structural Variations in Species B Adenovirus Fibers Impact CD46 Association. *J. Virol.* **82**, 7923–7931 (2008).
372. Howitt, J., Bewley, M. C., Graziano, V., Flanagan, J. M. & Freimuth, P. Structural Basis for Variation in Adenovirus Affinity for the Cellular Coxsackievirus and Adenovirus Receptor. *J. Biol. Chem.* **278**, 26208–26215 (2003).

373. Xia, D., Henry, L. J., Gerard, R. D. & Deisenhofer, J. Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution. *Struct. Lond. Engl.* **1993** *2*, 1259–1270 (1994).
374. Chen, C. Z., Gorham, R. D., Gaieb, Z. & Morikis, D. Electrostatic Interactions between Complement Regulator CD46(SCR1-2) and Adenovirus Ad11/Ad21 Fiber Protein Knob. *Molecular Biology International* <https://www.hindawi.com/journals/mbi/2015/967465/> (2015) doi:10.1155/2015/967465.
375. Trinh, H. V. *et al.* Avidity Binding of Human Adenovirus Serotypes 3 and 7 to the Membrane Cofactor CD46 Triggers Infection. *J. Virol.* **86**, 1623–1637 (2012).
376. Rainbow, A. J. & Castillo, J. E. Homologous recombination of adenovirus DNA in mammalian cells: enhanced recombination following UV-irradiation of the virus. *Mutat. Res.* **274**, 201–210 (1992).
377. Takeuchi, S. *et al.* Adenovirus Strains of Subgenus D Associated with Nosocomial Infection as New Etiological Agents of Epidemic Keratoconjunctivitis in Japan. *J. Clin. Microbiol.* **37**, 3392–3394 (1999).
378. Barouch, D. H. *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).
379. Cox, F. *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).
380. Dakin, R. S. *et al.* P238 Efficient gene transfer to human vascular cells in vitro and ex vivo using adenovirus serotype 49. *Cardiovasc. Res.* **103**, S42–S42 (2014).
381. Camacho, Z. T., Turner, M. A., Barry, M. A. & Weaver, E. A. CD46-Mediated Transduction of a Species D Adenovirus Vaccine Improves Mucosal Vaccine Efficacy. *Hum. Gene Ther.* **25**, 364–374 (2014).
- 382.377. CD46 Does Not Affect the In Vivo Biodistribution of Species D Adenoviruses, Ad26 or Ad48 in CD46 Transgenic Mice Following IV Delivery. *Mol. Ther.* **20**, S148 (2012).
383. Arnberg, N. *et al.* Adenovirus Type 37 Binds to Cell Surface Sialic Acid Through a Charge-Dependent Interaction. *Virology* **302**, 33–43 (2002).
384. Arnberg, N., Kidd, A. H., Edlund, K., Olfat, F. & Wadell, G. Initial Interactions of Subgenus D Adenoviruses with A549 Cellular Receptors: Sialic Acid versus α vIntegrins. *J. Virol.* **74**, 7691–7693 (2000).
385. Lei, Z. *et al.* Outbreaks of epidemic keratoconjunctivitis caused by human adenovirus type 8 in the Tibet Autonomous Region of China in 2016. *PLoS ONE* **12**, (2017).
386. Hage, E. *et al.* Molecular phylogeny of a novel human adenovirus type 8 strain causing a prolonged, multi-state keratoconjunctivitis epidemic in Germany. *Sci. Rep.* **7**, (2017).

387. de Jong, J. C. *et al.* Adenovirus 37: identification and characterization of a medically important new adenovirus type of subgroup D. *J. Med. Virol.* **7**, 105–118 (1981).
388. Ariga, T. *et al.* Five New Genome Types of Adenovirus Type 37 Caused Epidemic Keratoconjunctivitis in Sapporo, Japan, for More Than 10 Years. *J. Clin. Microbiol.* **43**, 726–732 (2005).
389. Aoki, K., Kitaichi, N., Hinokuma, R. & Ohno, S. Characteristics of Epidemic Keratoconjunctivitis due to Novel Adenovirus Types. **1**, 3 (2017).
390. Matsuura, K. *et al.* Human adenoviral type 54 keratoconjunctivitis accompanied by stellate keratitis and keratic precipitates: two cases. *BMC Ophthalmol.* **19**, 7 (2019).
391. Huang, G. *et al.* Outbreak of Epidemic Keratoconjunctivitis Caused by Human Adenovirus Type 56, China, 2012. *PLOS ONE* **9**, e110781 (2014).
392. Arcasoy, S. M. *et al.* MUC1 and Other Sialoglycoconjugates Inhibit Adenovirus-mediated Gene Transfer to Epithelial Cells. *Am. J. Respir. Cell Mol. Biol.* **17**, 422–435 (1997).
393. Esposito, S. *et al.* Epidemiology and Clinical Characteristics of Respiratory Infections Due to Adenovirus in Children Living in Milan, Italy, during 2013 and 2014. *PLOS ONE* **11**, e0152375 (2016).
394. Civljak, R. *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.

395. Blusch, J. H. *et al.* The novel early region 3 protein E3/49K is specifically expressed by adenoviruses of subgenus D: implications for epidemic keratoconjunctivitis and adenovirus evolution. *Virology* **296**, 94–106 (2002).
396. Feng, M., Chang, R. S.-M., Smith, T. R. & Snyder, J. C. Adenoviruses Isolated from Saudi Arabia. *Am. J. Trop. Med. Hyg.* **8**, 501–504 (1959).
397. Arnberg, N., Mei, Y. f & Wadell, G. Fiber genes of adenoviruses with tropism for the eye and the genital tract. *Virology* **227**, 239–244 (1997).
398. Seya, T., Turner, J. R. & Atkinson, J. P. Purification and characterization of a membrane protein (gp45-70) that is a cofactor for cleavage of C3b and C4b. *J. Exp. Med.* **163**, 837–855 (1986).
399. Seya, T., Hirano, A., Matsumoto, M., Nomura, M. & Ueda, S. Human membrane cofactor protein (MCP, CD46): multiple isoforms and functions. *Int. J. Biochem. Cell Biol.* **31**, 1255–1260 (1999).
400. Liszewski, M. K. & Atkinson, J. P. Membrane cofactor protein. *Curr. Top. Microbiol. Immunol.* **178**, 45–60 (1992).
401. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
402. Tissue expression of CXADR - Summary - The Human Protein Atlas. <https://www.proteinatlas.org/ENSG00000154639-CXADR/tissue>.
403. Tissue expression of DSG2 - Summary - The Human Protein Atlas. <https://www.proteinatlas.org/ENSG00000046604-DSG2/tissue>.
404. Schäfer, S., Koch, P. J. & Franke, W. W. Identification of the ubiquitous human desmoglein, Dsg2, and the expression catalogue of the desmoglein subfamily of desmosomal cadherins. *Exp. Cell Res.* **211**, 391–399 (1994).

405. Varki, A. Sialic acids in human health and disease. *Trends Mol. Med.* **14**, 351–360 (2008).
406. Webster, B., Ott, M. & Greene, W. C. Evasion of superinfection exclusion and elimination of primary viral RNA by an adapted strain of hepatitis C virus. *J. Virol.* **87**, 13354–13369 (2013).
407. Sun, J. & Brooke, C. B. Influenza A Virus Superinfection Potential Is Regulated by Viral Genomic Heterogeneity. *mBio* **9**, e01761-18 (2018).
408. Liu, S. *et al.* Tight Junction Proteins Claudin-1 and Occludin Control Hepatitis C Virus Entry and Are Downregulated during Infection To Prevent Superinfection. *J. Virol.* **83**, 2011–2014 (2009).
409. Berngruber, T. W., Weissing, F. J. & Gandon, S. Inhibition of Superinfection and the Evolution of Viral Latency. *J. Virol.* **84**, 10200–10208 (2010).
410. Folimonova, S. Y. Superinfection Exclusion Is an Active Virus-Controlled Function That Requires a Specific Viral Protein. *J. Virol.* **86**, 5554–5561 (2012).
411. Ismail, A. M. *et al.* Genomic analysis of a large set of currently—and historically—important human adenovirus pathogens. *Emerg. Microbes Infect.* **7**, (2018).
412. Li, P. *et al.* Circulation of HAdV-41 with diverse genome types and recombination in acute gastroenteritis among children in Shanghai. *Sci. Rep.* **7**, 3548 (2017).

413. Human mastadenovirus D strain
human/DEU/HEIM_00083/1986/100[P100H17F30], complete genome.
(2019).
414. Human mastadenovirus D strain
human/DEU/HEIM_00096/1984/102[P102H38F30], complete genome.
(2019).
415. Human mastadenovirus D strain
human/DEU/HEIM_00091/1986/103[P103H33F30], complete genome.
(2019).
416. James, L. *et al.* Outbreak of Human Adenovirus Type 3 Infection in a Pediatric Long-Term Care Facility—Illinois, 2005. *Clin. Infect. Dis.* **45**, 416–420 (2007).
417. Wadell, G., Varsányi, T. M., Lord, A. & Sutton, R. N. P. EPIDEMIC OUTBREAKS OF ADENOVIRUS 7 WITH SPECIAL REFERENCE TO THE PATHOGENICITY OF ADENOVIRUS GENOME TYPE 7b. *Am. J. Epidemiol.* **112**, 619–628 (1980).
418. Ryan, M. A. K. *et al.* Large Epidemic of Respiratory Illness Due to Adenovirus Types 7 and 3 in Healthy Young Adults. *Clin. Infect. Dis.* **34**, 577–582 (2002).
419. Comer, F. I. & Hart, G. W. O-Glycosylation of Nuclear and Cytosolic Proteins DYNAMIC INTERPLAY BETWEEN O-GlcNAc AND O-PHOSPHATE. *J. Biol. Chem.* **275**, 29179–29182 (2000).
420. King, S. L. *et al.* Characterizing the O-glycosylation landscape of human plasma, platelets, and endothelial cells. *Blood Adv.* **1**, 429–442 (2017).

421. Munkley, J. The glycosylation landscape of pancreatic cancer. *Oncol. Lett.* **17**, 2569–2575 (2019).
422. Gamblin, S. J. & Skehel, J. J. Influenza Hemagglutinin and Neuraminidase Membrane Glycoproteins. *J. Biol. Chem.* **285**, 28403–28409 (2010).
423. Kosik, I. & Yewdell, J. W. Influenza Hemagglutinin and Neuraminidase: Yin–Yang Proteins Coevolving to Thwart Immunity. *Viruses* **11**, (2019).
424. Lofgren, E., Fefferman, N. H., Naumov, Y. N., Gorski, J. & Naumova, E. N. Influenza Seasonality: Underlying Causes and Modeling Theories. *J. Virol.* **81**, 5429–5436 (2007).
425. Ismail, A. M. *et al.* Adenoviromics: Mining the Human Adenovirus Species D Genome. *Front. Microbiol.* **9**, (2018).
426. Thorner, A. R. *et al.* Age Dependence of Adenovirus-Specific Neutralizing Antibody Titers in Individuals from Sub-Saharan Africa. *J. Clin. Microbiol.* **44**, 3781–3783 (2006).
427. Curlin, M. E. *et al.* Frequent Detection of Human Adenovirus from the Lower Gastrointestinal Tract in Men Who Have Sex with Men. *PLOS ONE* **5**, e11321 (2010).
428. Byrd-Leotis, L., Cummings, R. D. & Steinhauer, D. A. The Interplay between the Host Receptor and Influenza Virus Hemagglutinin and Neuraminidase. *Int. J. Mol. Sci.* **18**, (2017).
429. Yetter, R. A., Lehrer, S., Ramphal, R. & Small, P. A. Outcome of Influenza Infection: Effect of Site of Initial Infection and Heterotypic Immunity. *Infect. Immun.* **29**, 654–662 (1980).

430. Lai, J. C. C., Karunaratna, H. M. T. K., Wong, H. H., Peiris, J. S. M. & Nicholls, J. M. Neuraminidase activity and specificity of influenza A virus are influenced by haemagglutinin-receptor binding. *Emerg. Microbes Infect.* **8**, 327–338 (2019).
431. Pfortmueller, C. A. *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).
432. Chang, S.-Y. *et al.* A community-derived outbreak of adenovirus type 3 in children in Taiwan between 2004 and 2005. *J. Med. Virol.* **80**, 102–112 (2008).
433. Alissa Alkhalaf, M., Al Qurashi, Y. M. A., Guiver, M. & Cooper, R. J. Genome Sequences of Three Species D Adenoviruses Isolated from AIDS Patients. *Genome Announc.* **2**, (2014).
434. Rosen, L., Baron, S. & Bell, J. A. Four Newly Recognized Adenoviruses. *Proc. Soc. Exp. Biol. Med.* **107**, 434–437 (1961).
435. Human adenovirus 26 gene for fiber protein, partial cds, strain: BP-2. (2016).
436. Ramya, T. N. C. *et al.* In Situ trans Ligands of CD22 Identified by Glycan-Protein Photocross-linking-enabled Proteomics. *Mol. Cell. Proteomics MCP* **9**, 1339–1351 (2010).
437. Drake, R. R. *et al.* MALDI mass spectrometry imaging of N-linked glycans in cancer tissues. *Adv. Cancer Res.* **134**, 85–116 (2017).

438. Nita-Lazar, M. *et al.* Desialylation of airway epithelial cells during influenza virus infection enhances pneumococcal adhesion via galectin binding. *Mol. Immunol.* **65**, 1–16 (2015).
439. Varki, A. Letter to the Glyco-Forum: Since there are PAMPs and DAMPs, there must be SAMPs? Glycan “self-associated molecular patterns” dampen innate immunity, but pathogens can mimic them. *Glycobiology* **21**, 1121–1124 (2011).
440. Chang, Y.-C., Uchiyama, S., Varki, A. & Nizet, V. Leukocyte Inflammatory Responses Provoked by Pneumococcal Sialidase. *mBio* **3**, e00220-11 (2012).
441. Stamatos, N. M., Curreli, S., Zella, D. & Cross, A. S. Desialylation of glycoconjugates on the surface of monocytes activates the extracellular signal-related kinases ERK 1/2 and results in enhanced production of specific cytokines. *J. Leukoc. Biol.* **75**, 307–313 (2004).
442. Pappu, B. P. & Shrikant, P. A. Alteration of Cell Surface Sialylation Regulates Antigen-Induced Naive CD8+ T Cell Responses. *J. Immunol.* **173**, 275–284 (2004).
443. Thorner, A. R. *et al.* Immunogenicity of Heterologous Recombinant Adenovirus Prime-Boost Vaccine Regimens Is Enhanced by Circumventing Vector Cross-Reactivity. *J. Virol.* **80**, 12009–12016 (2006).
444. Saxena, M. & Bhardwaj, N. Re-Emergence of Dendritic Cell Vaccines for Cancer Treatment. *Trends Cancer* **4**, 119–137 (2018).
445. Mastelic-Gavillet, B., Balint, K., Boudousquie, C., Gannon, P. O. & Kandalaft, L. E. Personalized Dendritic Cell Vaccines—Recent

Breakthroughs and Encouraging Clinical Results. *Front. Immunol.* **10**,
(2019).

Publications

Baker, A.T. et al. 2018. *Designer Oncolytic Adenovirus: Coming of Age.* *Cancers* 10(6), p. 201. doi: 10.3390/cancers10060201.

Baker, A.T. et al. 2019a. *Diversity within the adenovirus fiber knob hypervariable loops influences primary receptor interactions.* *Nature Communications* 10(1), p. 741. doi: 10.1038/s41467-019-08599-y.

Baker, A.T. et al. 2019b. *Human adenovirus type 26 uses sialic acid-bearing glycans as a primary cell entry receptor.* *Science Advances* 5(9), p. eaax3567. doi: 10.1126/sciadv.aax3567.

Qin, Z. et al. 2013. The Fish Pathogen Yersinia ruckeri Produces Holomycin and Uses an RNA Methyltransferase for Self-resistance. Journal of Biological Chemistry 288(21), pp. 14688–14697. doi: 10.1074/jbc.M112.448415.

Uusi-Kerttula, H. et al. 2018. Ad5NULL-A20 - a tropism-modified, $\alpha\beta 6$ integrin-selective oncolytic adenovirus for epithelial ovarian cancer therapies. Clinical Cancer Research , p. clincanres.1089.2018. doi: 10.1158/1078-0432.CCR-18-1089.

ARTICLE

<https://doi.org/10.1038/s41467-019-08599-y>

OPEN

Diversity within the adenovirus fiber knob hypervariable loops influences primary receptor interactions

Alexander T. Baker¹, Alexander Greenshields-Watson², Lynda Coughlan ³, James A. Davies¹, Hanni Uusi-Kerttula¹, David K. Cole^{2,4}, Pierre J. Rizkallah² & Alan L. Parker ¹

Adenovirus based vectors are of increasing importance for wide ranging therapeutic applications. As vaccines, vectors derived from human adenovirus species D serotypes 26 and 48 (HAdV-D26/48) are demonstrating promising efficacy as protective platforms against infectious diseases. Significant clinical progress has been made, yet definitive studies underpinning mechanisms of entry, infection, and receptor usage are currently lacking. Here, we perform structural and biological analysis of the receptor binding fiber-knob protein of HAdV-D26/48, reporting crystal structures, and modelling putative interactions with two previously suggested attachment receptors, CD46 and Coxsackie and Adenovirus Receptor (CAR). We provide evidence of a low affinity interaction with CAR, with modelling suggesting affinity is attenuated through extended, semi-flexible loop structures, providing steric hindrance. Conversely, *in silico* and *in vitro* experiments are unable to provide evidence of interaction between HAdV-D26/48 fiber-knob with CD46, or with Desmoglein 2. Our findings provide insight into the cell-virus interactions of HAdV-D26/48, with important implications for the design and engineering of optimised Ad-based therapeutics.

¹Division of Cancer and Genetics, School of Medicine, Cardiff University, Cardiff CF14 4XN, UK. ²Division of Infection and Immunity, School of Medicine, Cardiff University, Cardiff CF14 4XN, UK. ³Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029-6574, USA. ⁴Present address: Immunocore Ltd., Abingdon OX14 4RY Oxon, UK. Correspondence and requests for materials should be addressed to A.L.P. (email: ParkerAL@Cardiff.ac.uk)

Adenoviruses are increasingly important vectors for wide-ranging therapeutic interventions, from gene delivery and oncolytic agents to platforms for vaccine applications^{1–3}. As vaccine vectors, their use clinically has been popularised by their excellent safety profile coupled with their ability to induce robust cellular and humoral immunogenicity in humans⁴. Phylogenetically, the human adenoviruses (HAdV's) are diverse, subdivided across seven species, A–G⁵, based classically on serological cross-reactivity, receptor usage, haemagglutination properties and, more recently, phylogenetic sequence similarity^{6,7}.

Most experimental and clinical studies have focussed on the well-studied species C adenovirus, HAdV-C5. Although potentially immunogenic, the efficacy of vaccines based on HAdV-C5 appears hampered by high seroprevalence rates in humans, and enthusiasm for their use as clinical vaccine platforms has been dampened by the well-publicised failure of the MERCK sponsored STEP vaccine trial. This trial, to evaluate an HAdV-C5-based HIV vaccine encoding HIV gag/pol/nef antigens, was abandoned due to apparent lack of efficacy upon 1st term analysis. The study also identified a non-significant trend towards increased HIV acquisition in a specific high-risk, uncircumcised subset of patients who also had high levels of baseline pre-existing neutralising antibodies (NAbs) to HAdV-C5^{8,9}. As a result, attention has switched from HAdV-C5-based vectors towards the development of alternative adenoviral serotypes with lower rates of pre-existing immunity. Most notably, vectors under development include those based on species D serotypes including HAdV-D26, which has entered Phase-III clinical trials as an Ebola vaccine and recently reported promising immunogenicity in an HIV trial. Chimeric vectors utilising the hexon hyper variable regions (HVRs) of HAdV-D48 which have undergone Phase-I evaluation as an HIV vaccine^{3,10–12}. However, despite extensive clinical advances using these vaccine vectors we possess very limited knowledge of their basic biology, particularly with regards to the determinants underpinning their tropism, mechanisms of cellular entry, and receptor usage. In this study, we address these shortcomings through analysis of adenoviral diversity in the context of their receptor binding, fiber proteins. Whilst adenoviruses are historically divided into seven species, A–G, this may underestimate their diversity^{13–15}. Phylogenetic examination of 56 human adenovirus fiber proteins from different species shows deviation from the taxonomy expected based upon whole genome taxonomy, likely due to recombination events as seen in other adenoviral proteins¹⁴. Here, we have sought to generate high-resolution crystallographic structures of the cellular interacting fiber-knob domains of species D adenoviruses HAdV-D26 and HAdV-D48. The fiber-knob is the receptor interacting domain of the fiber protein, one of three major capsid proteins along with the hexon and penton base, as shown schematically in Fig. 1a.

In this study, we employ an integrative workflow utilising X-ray crystallography, *in silico* modelling, and *in vitro* assays to dissect previous findings^{16,17} suggesting interactions by HAdV-D26 and HAdV-D48 with Coxsackie and Adenovirus Receptor (CAR)^{5,18} and CD46 (Membrane Cofactor Protein, MCP)^{16,17,19–21}. Utilising surface plasmon resonance (SPR), we also investigate the potential for HAdV-D26 and HAdV-D48 to interact with Desmoglein 2. Our findings shed new light on the cell–virus interactions of adenovirus and have potential implications for the design and engineering of optimised HAdV-based therapeutics, both for vaccine applications and oncolytic development, allowing us to minimise off-target or undesirable interactions *in vivo*.

Results

Genetic variability in adenovirus fiber-knob protein. We generated phylogenetic trees of human adenovirus serotypes 1–56 which revealed greater diversity in the fiber-knob domain (Fig. 1b) than might be expected based upon the taxonomy of the whole virus (Fig. 1c). These phylogenetic trees have been condensed to 70% bootstrap confidence (500 bootstrap replicates) to exclude poorly supported nodes and display the projected diversity. A full dendrogram showing to-scale branches is provided in Supplementary Figure 1. In both phylogenetic trees, the adenoviruses divide into seven clades corresponding to the seven adenoviral species, A–G. However, the tree based upon the fiber-knob domain (Fig. 1b) shows the species D adenoviruses forming a greater number of sub-groups than in the whole genome tree suggesting greater diversity in the receptors of this species than might be expected when comparing serotypes. This may be the result of recombination events, as reported previously for other adenoviral proteins^{13,15,22–24}. The opposite is observed in species B adenoviruses, where simpler groupings are seen when analysing fiber-knob domain alignment than by whole genome analysis. When divided into sub-species based on whole genome, they divide into species B1 and B2, but when the tree is generated based on fiber-knob alone the species B viruses do not divide into similar groups. The significance of this fiber-knob diversity is unclear, but it has previously been suggested that the species B1/B2 designation may more closely represent the tissue tropism, than receptor usage^{25,26}.

We next calculated the amino acid variability at each position in the aligned adenoviral fiber-knob sequences, which revealed regions of broad conservation corresponding to β -strands which make up the main fold of the fiber-knob trimer (Fig. 1d). The positions corresponding to the β -strands of HAdV-C5, as originally reported by Xia et al.²⁷, are shown by arrows. Comparison shows that the more N-terminal (A, B, and C) and C-terminal (I and J) β -strands have greater homology across the adenoviral species than the other strands. This may relate to the intervening loops between the less tightly conserved β -strands (D, E, F, G, H) being more apical, a region which is often involved in receptor interactions^{26,28–32}.

Structural analysis of HAdV-D26 and HAdV-D48 fiber-knob.

To investigate diversity within the species D adenoviruses fiber-knob protein, recombinant, 6-His-tagged fiber-knob protein from HAdV-D26 and HAdV-D48 (hereafter referred to as HAdV-D26K and HAdV-D48K) were generated, purified, and used to determine X-ray crystallographic structures of HAdV-D26 (PDB: 6FJN) and HAdV-D48 (PDB: 6FJQ) at resolutions of 0.97 and 2.91 Å, respectively (Fig. 2a–d). Table 1 shows the data collection and refinement statistics for the crystallographic structures generated in this study.

The monomers (red, green, blue) form an anti-parallel β -barrel, typical of adenoviral fiber-knob protein, as described by Xia et al. (PDB 1KNB)²⁷. Each monomer interacts with two neighbouring copies to form a homotrimer with 3-fold symmetry (Fig. 2b, d) and a highly stable interface (Supplementary Figure 2A). Stability analysis using PISA (Protein Interactions, Surfaces and Assemblies) software calculates the HAdV-D26 and HAdV-D48 trimers to have >20% lower interface energy than that of HAdV-C5, indicative of a more stable interaction (Supplementary Figure 2A) between the monomers of the HAdV-D26 and HAdV-D48 fiber-knobs.

As with the pan species analysis (Fig. 1d), variability of the aligned species D fiber-knobs (Fig. 2e) confirmed that the β -strands comprising the hydrophobic core of the trimers are

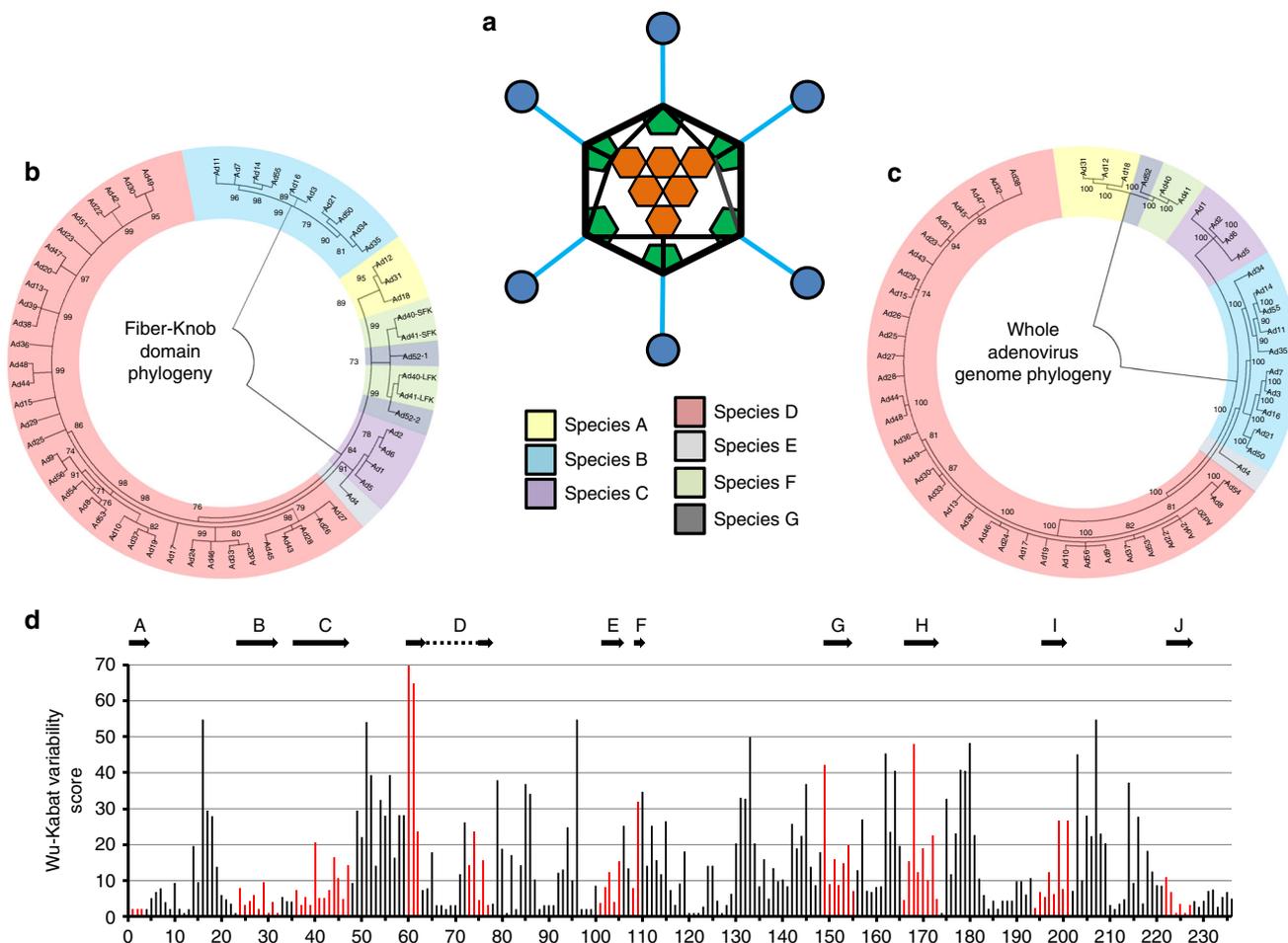


Fig. 1 Phylogenetic analysis of adenoviruses mapped by whole genome and fiber-knob domain. A diagrammatic representation of the adenoviral major capsid proteins shows the icosahedral capsid structure with the Hexon (orange) comprising the facets, pentons (green) at the vertices, from which the fiber proteins (fiber-shaft in light blue, fiber-knob in dark blue) protrude (a). Condensed maximum likelihood trees (percentage confidence shown by numbers next to nodes) were generated from alignments of fiber-knob domain amino acid sequences of adenoviruses 1–56 (b) or whole genome NT sequences (c). Adenoviruses divide into 7 subspecies, as denoted in the key, regardless of alignment used, but the species D adenoviruses divide into additional sub-species when determined by fiber-knob alignment, for readability simple nomenclature is used, all are human adenovirus. Numbers next to nodes denote confidence. Wu-Kabat variability analysis of the Clustal Omega aligned fiber-knob domains amino acid sequences of adenoviruses 1–56 (d) reveals regions of low amino acid variability corresponding to beta-sheets. The locations of HAdV-C5 β -strands, as described by Xia et al. (1994)²⁷, are aligned to the structure and shown by arrows, the corresponding positions are coloured in red

highly conserved, with β -strands demonstrating a high degree of overlap in both spatial position and sequence variability (Supplementary Figure 2B).

Fiber-knob loops are stabilised by inter-loop interactions.

Particularly relevant to this study are the DG, GH, HI, and IJ loops, linking the indicated strands corresponding to those in the originally reported HAdV-C5 fiber-knob loops (Supplementary Figure 2B)²⁷. These loops have previously been shown to be critical in engagement of CD46 for Ad11, 35, and 21^{33–36}. Alignment of these loops with the corresponding loops of HAdV-C5 (species C, CAR interacting), HAdV-B35 (species B, CD46 interacting), Ad11 (species B, CD46 interacting), and Ad37 (species D, CAR and Sialic acid/GD1a glycan interacting) reveals different topologies in these critical receptor interacting regions (Fig. 3a)^{35–38}. The HI loops of HAdV-D26K and HAdV-D48K are most homologous to those of HAdV-B35K and HAdV-D37K respectively in terms of amino acid sequence identity (Supplementary Figure 2B) and spatial alignment. The HAdV-D26K DG loop is most homologous to HAdV-B35K but incorporates a

three amino acid insertion, while HAdV-D48K DG loop displays a differing and unique topology. The GH and IJ loops of HAdV-C5K, HAdV-D26K, HAdV-D37K, and HAdV-D48K demonstrate similar spatial arrangements (though the similarity does not extend to the sequence identity) but differ from the CD46 utilising HAd-B11K and HAdV-B35K.

The high-resolution structures obtained allowed us to robustly characterise the loops, seen in the electron density maps (Fig. 3b, c). To assess loop flexibility and mobility, we assessed the B-factors (also known as temperature factor), a measure of the confidence in the position of an atom which can be used to infer flexibility. By assessing the B-factors, the relative flexibility of the moieties of interest can be inferred. While the apical domains of some loops demonstrated increased B-factors relative to the rest of the molecule, the loops’ B-factors are surprisingly low (Fig. 3b, c), suggesting that they may exhibit limited flexibility.

To investigate whether the different loop conformations were the product of flexibility, or restricted mobility we investigated the inter-loop interactions in the HAdV-D26K and HAdV-D48K structures. This analysis shows that the GH loop of HAdV-D26K (like those of HAdV-C5K, HAdV-D37K, and HAdV-D48K) does

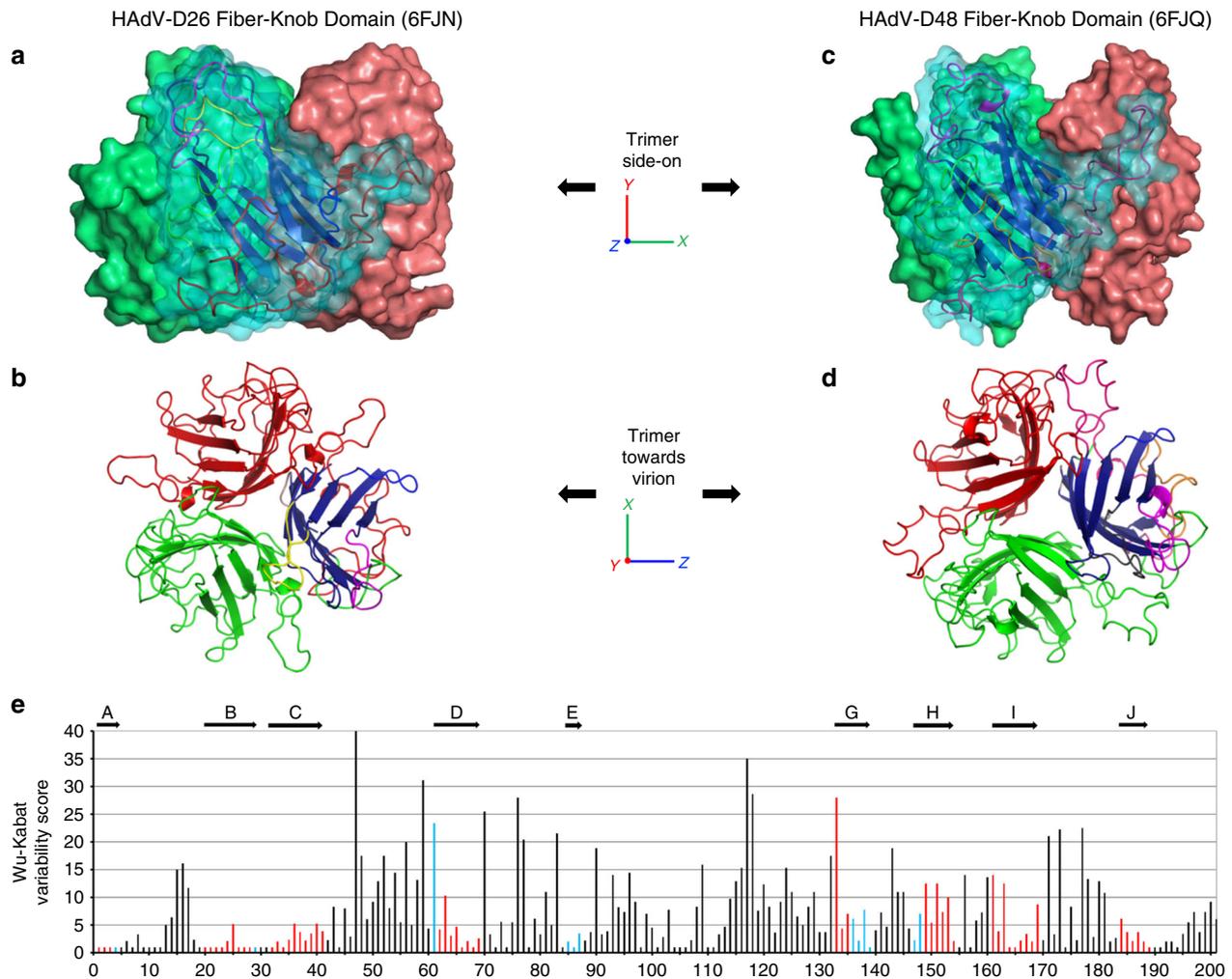


Fig. 2 Overview of the HAdV-D26 and HAdV-D48 fiber-knob protein structures. The surface representation of the trimeric HAdV-D26K (PDB 6FJN) biological assembly is shown side-on with the cartoon representation shown for the nearest monomer (**a**) and the top-down view of the same HAdV-D26K trimer, as it would appear looking towards the virion, is seen as a cartoon representation (**b**), with each monomer coloured in red, green, or blue with the hypervariable loops extending between the β -strands (dark blue) coloured as follows: AB—green, BC—purple, CD—brown, DE—orange, DG—red, EG—pink, GH—purple, HI—yellow, IJ—light blue. The HAdV-D48K (PDB 6FJQ) trimer is shown similarly (**c**, **d**). The Wu-Kabat variability plot of the fiber-knob domains of species D adenoviruses shows regions of low variability (**e**) with the locations of the HAdV-D48K β -strands shown by arrows above the graph, and the positions coloured blue, or red when the position is a β -strand in both HAdV-D26K and HAdV-D48K

not extend directly away from the G and H β -sheets, but forms a β -hairpin (Fig. 3a, b) that is maintained by seven polar contacts within the neighbouring IJ loop which restrict the loops' orientation (Fig. 4a, Supplementary Figure 3A). Polar contacts were also observed at the apex of several loops, notably the GH and CD loops (Fig. 4a, b). The IJ loops form fewer intramolecular polar contacts but are stabilised by interactions with the adjacent CD and DG loops (Supplementary Figure 3C, D). These interactions retain the apical residues in a stable conformation, rather than leaving the side chains fully labile.

The B-factors of the HAdV-D48K DG loop were observed to be polarised about the hairpin, with the outer face of the loop having higher B-factors compared to the inner face (Fig. 4c). This is likely the result of polar contacts formed between Ser-307, Gln-308, Ala-309, and Leu-304 with Asp-359 and Gln-357 of the opposing monomer stabilising the conformation of the DG loop. The proline-rich nature of this loop provides further rigidity (Supplementary Figure 2D).

Crystal contacts did not reveal any specific interactions between these DG-loops and neighbouring non-trimer copies. We calculated the energy of interaction to be below the

background threshold ($< -3.0 \text{ kcal mol}^{-1}$) for all loops except DG. The DG-loop of HAdV-D26K is calculated to have interaction energy of $-6.5 \text{ kcal mol}^{-1}$ in two separate stretches of this exceptionally long loop (Supplementary Figure 4). Importantly, no strong contacts are found within the inter-monomer cleft.

Based on this analysis of the inter and intra-loop bonds, we suggest that these adenoviral loops may not be fully flexible variable regions, but organised receptor engagement motifs with carefully evolved structures. This has direct implications for receptor engagement of these viruses, as the loops govern previously characterised interactions with CAR and CD46, and are directly involved in their pathogenicity^{25,33,39}.

In silico evaluation of HAdV-D26/48K interaction with CAR.

Both CD46 and CAR have been proposed as primary attachment receptors for HAdV-D26 and HAdV-D48^{16,17}. Previously generated crystal structures of HAdV-B11K in complex with full-length CD46 (PDB: 3O8E), and HAd-D37K in complex with CAR-D1 domain (PDB: 2J12) reveal the loops to be essential to

Table 1 Data collection and refinement statistics (molecular replacement)

	HAdV-D26K	HAdV-C5K	HAdV-D48K
Data collection			
Space group	P 2 ₁ 3	P 2 ₁ 2 ₁ 2	P 4 ₃ 2
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	86.01, 86.01, 86.01	102.16, 102.44, 77.01	145.18, 145.18, 145.18
α , β , γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	0.97–60.82 (0.97–1.00)	1.49–61.56 (1.49–1.53)	2.91–83.82 (2.91–2.99)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.043 (0.745)	0.134 (1.838)	0.125 (302.6)
<i>I</i> / σ	27.3 (0.7)	7.1 (0.7)	22.2 (1.7)
Completeness (%)	94.9 (43.9)	99.8 (99.9)	100.0 (100)
Redundancy	16.7 (1.6)	6.6 (6.3)	41.2 (41.4)
Refinement			
Resolution (Å)	0.97–60.82	1.49–61.56	2.91–83.82
No. reflections	112,612	125,479	11,371
<i>R</i> _{work} / <i>R</i> _{free}	18.2/19.5	21.1/23.3	20.1/29.1
No. atoms	1811	4825	3117
Protein	1579	4395	3091
Ligand/ion	8	21	20
Water	224	409	6
<i>B</i> -factors	16.0	34.0	87.0
Protein	15.8	33.6	94.2
Ligand/ion	29.3	35.7	129.5
Water	23.9	42.9	61.0
R.M.S. deviations			
Bond lengths (Å)	0.025	0.011	0.019
Bond angles (°)	2.339	1.534	2.293

One crystal was used for each dataset
Values in parentheses are for highest-resolution shell

receptor interactions⁴⁰. To investigate the ability of HAdV-D26K and HAdV-D48K to bind these receptors we generated homology models by alignment of the new HAdV-D26K and HAdV-D48K fiber-knob structures modelled, using the existing fiber-knobs in complex with the receptor of interest, and performed energy minimisation to optimise the conformation to achieve the lowest possible energy interface with which to analyse the interaction. We performed similar experiments with the well-described CAR and CD46 binding fiber-knob proteins of HAdV-C5 (PDB: 6HCN) and HAdV-B35 (PDB: 2QLK), respectively, as controls.

Modelling of HAdV-D26K and HAdV-D48K in complex with the CAR-D1 domain revealed a region of high homology with the CAR utilising HAdV-C5 fiber-knob, hereafter termed the α -interface (Fig. 5a). Sequence alignment with HAdV-C5K shows that many of the residues previously shown to be critical for CAR interaction in HAdV-C5K are conserved in HAdV-D26K and HAdV-D48K (Fig. 5b), including Ser-408, Pro-409, and Tyr-376³⁷. The same is true of residues predicted to interact with CAR directly, such as Lys-417 (number is for HAdV-C5K). The residues predicted to form direct CAR binding interactions for HAdV-C5K, HAdV-D26K, and HAdV-D48K are pictured in complex with the maximum spatial occupancy of the energy minimised CAR-D1 (Fig. 5c). The high levels of homology are seen to extend to the proteins' fold as well as the linear sequence.

Binding energies were calculated between the modelled fiber-knob proteins and CAR, restricting the calculation to only the α -interface to best model the conserved region. For the modelled complexes, a stable α -interface was predicted for all complexes modelled, albeit weaker for the known non-CAR utilising HAdV-B35K (Fig. 5d) which has lower sequence conservation with HAdV-C5K. However, the interaction is complicated by a second CAR interface, termed the β -interface (Fig. 6a). The loops

forming the β -interface with CAR-D1 differ between HAdV-C5, HAdV-D26, and HAdV-D48 fiber-knob (Fig. 6b). The shorter HAdV-C5K DG loop does not clash with the CAR-D1 surface, whereas the extended HAdV-D26K DG loop forms a partial steric clash, with surface seen to clash with the aligned CAR-D1, and HAdV-D48K DG loop is seen to form an even larger steric clash. Whilst the longer loop of HAdV-D26K is expected to be more flexible than that of HAdV-C5K, the HAdV-D48K DG loop is surprisingly stable due to the characteristics described (Fig. 4c).

Biological evaluation of HAdV-D26/48K interaction with CAR. Our modelling studies indicate that the longer, inflexible DG loop of HAdV-D48K would be likely to sterically hinder the HAdV-D48K: CAR interaction at the β -interface to a greater extent than the more modest inhibition of the smaller and more labile loop of HAdV-D26K, which in turn would exhibit more inhibition of CAR binding than that of HAdV-C5K, where no steric inhibition is observed. Competition inhibition assays using recombinant fiber-knob protein to inhibit antibody binding to CAR receptor in CHO-CAR cells (which express CAR, while the parental cell line (CHO-K1) is established to be non-permissive to adenovirus infection) support our observations (Fig. 6c). The IC₅₀ (the concentration of protein required to inhibit 50% of antibody binding) of HAdV-C5K is 7.0 ng/10⁵ cells, while HAdV-D26K and HAdV-D48K demonstrate IC₅₀ values 15.7 and 480 times higher at 0.110 μ g/10⁵ cells and 3.359 μ g/10⁵ cells, respectively, reflecting their reduced ability to engage CAR.

SPR analysis indicates that HAdV-C5K binds strongly to CAR (Fig. 6d) with a *K*_D of 0.76 nM. HAdV-D26K and HAdV-D48K have lower overall affinities for CAR (Fig. 6e). While the *K*_{off} of the three fiber knob proteins (Fig. 6e) are similar, the *K*_{on} is fastest for HAdV-C5K, with HAdV-D26K *K*_{on} being slower, and HAdV-D48K exhibiting the slowest *K*_{on}. This shows that the *K*_{on}—the ability to form the initial interaction with the receptor—is the major limiting factor in the fiber knobs overall affinity for CAR.

In silico evaluation of HAdV-D26/48K interaction with CD46.

A similar approach was taken to model HAdV-D26K and HAdV-D48K in complex with CD46. Alignments with the previously published HAdV-B11K-CD46 complex were generated and energy minimised to obtain the lowest energy state of the complex^{34,41}. This interface utilises loops DG, GH, HI, and IJ to form a network of polar interactions with the CD46 Sc1 and Sc2 domains (Fig. 7a)³⁵.

When HAdV-B35K was modelled in complex with CD46, a network of polar contacts between HAdV-B35K and CD46 was predicted (Fig. 7b) similar to that observed in the HAdV-B11K-CD46 complex crystal structure, PDB: 3O8E (Fig. 7a)³⁴. Previously, HAdV-B35K residues Phe-243, Arg-244, Tyr-260, Arg-279, Ser-282, and Glu-302 (underlined in Fig. 7c) have been implicated as key contact residues for CD46 interaction, and are conserved in HAdV-B11K (highlighted in blue, Fig. 7c)³⁶. Our modelling suggests that conversely, HAdV-D26K and HAdV-D48K are predicted to form very few polar contacts with CD46 with just two contacts predicted for HAdV-D26K (Fig. 7d) and three predicted for HAdV-D48K (Fig. 7e). Furthermore, they do not share any of the critical CD46 binding residues which have been reported previously (underlined, Fig. 7c) for HAdV-B11K and HAdV-B35K, or any of the predicted interacting residues (blue highlight, Fig. 7c).

We again employed PISA to calculate the binding energy of the various modelled and energy minimised fiber-knob CD46 complexes (Fig. 8a). HAdV-B11K, the strongest known CD46 binding adenovirus³³, was predicted to have the lowest binding

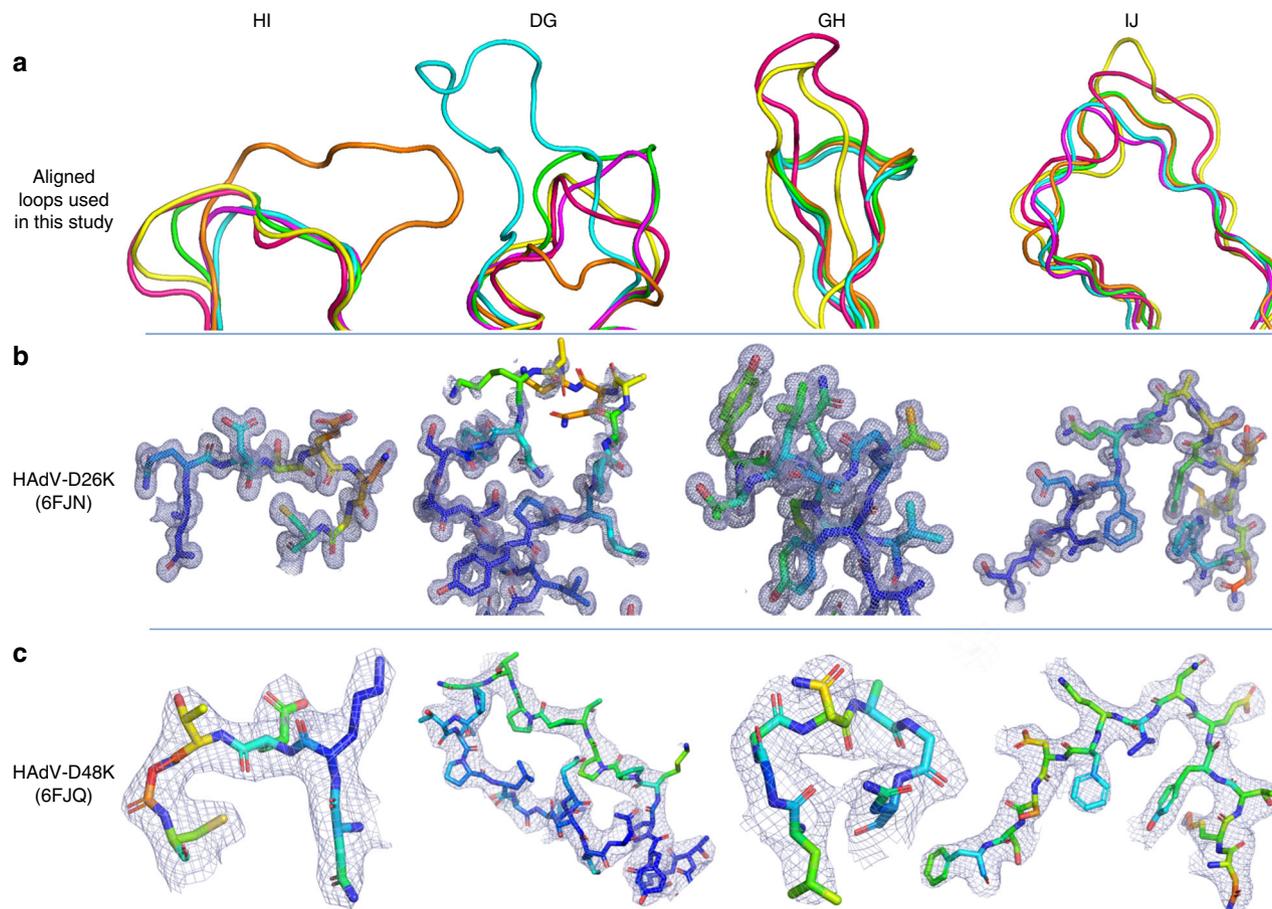


Fig. 3 Comparison of HI, DG, GH, and IJ loops of adenoviruses used in this study. **a** The hypervariable loops of HAdV-D26K (green) and HAdV-D48K (cyan) relevant to this study (HI, DG, GH, and IJ) are shown in the context of the control virus fiber-knob domains, HAdV-C5K (orange), Ad11K (yellow), HAdV-B35K (pink), and Ad37K (purple). The electron density achieved in the loops of HAdV-D26K (**b**) and HAdV-D48K (**c**) are shown as mesh. The fitted residues are seen as stick representations, with oxygen and nitrogen atoms coloured red and blue, respectively and other atoms coloured according to their relative B-factors with warmer colours indicating higher B-factor values

energy reflecting its high stability interface, with HAdV-B35 demonstrating a similar but slightly reduced binding energy. Conversely, HAdV-D26K and HAdV-D48K are predicted to have lower binding energies, similar to that which may be expected for random proteins passing in solution, indicating that any interaction between either the HAdV-D26K, or HAdV-D48K with CD46 is unlikely⁴².

While still low compared to the known CD46 interacting HAdV-B11K and HAdV-B35K binding energies, that for HAdV-C5K was higher than expected for a known non-CD46 interacting adenovirus (Fig. 8a). Inspection of the model shows that this is due to the close proximity of the large HAdV-C5 HI loop to CD46 (Fig. 8b). The residues involved in the predicted interaction are not conserved in any known CD46 interface and suggesting these are random interactions. Furthermore, interaction between the DG loop and CD46 is integral to known CD46 binding interfaces and is prevented by the HAdV-C5K HI loop laying between them (Fig. 8b).

Biological evaluation of HAdV-D26/48K interaction with CD46. Antibody competition inhibition assays in CHO-BC1 cells (CHO cells transduced to express the BC1 isoform of CD46) were used to test the predictions made by modelling (Fig. 8c). These data confirm that recombinant HAdV-D26K and HAdV-D48K proteins are incapable of inhibiting antibody binding to CD46 at

any concentration tested (up to 2 ng/cell), whilst the well-defined CD46 interacting HAdV-B35K demonstrates strong inhibition, with a calculated IC_{50} of $0.003 \mu\text{g}/10^5$ cells.

SPR analysis of the interaction between recombinant fiber-knob protein with CD46 confirms these findings. The known CD46 utilising HAdV-B35K is seen to bind CD46 even at low concentration, while HAdV-D48K shows no interaction (Fig. 8d). HAdV-D26K shows a very low affinity interaction with CD46, however the kinetics are extremely fast making it impossible to measure an accurate $K_{\text{On/Off}}$ at any of the concentrations measured, suggesting an unstable interface. The calculated K_D for HAdV-D26K is seen to be more than 1.5×10^3 times lower than that of HAdV-B35K (Fig. 8e).

In silico evaluation of HAdV-D26/48K interaction with DSG-2. The third major protein receptor for human adenoviruses is Desmoglein 2 (DSG2), shown to enable infection by HAdV-B3, B7, B11, and B14^{43,44}. Whilst we have not been able to model the interaction of HAdV-D26 or D48 with DSG2, due to the lack of an available high-resolution complexed structure at the time of writing, we investigated the interaction by SPR analysis. HAdV-B3 is the best studied DSG2 binding adenovirus and showed binding in the μM range when tested by SPR (Fig. 9a), however, no binding was observed when the same experiment was run with HAdV-D26K or HAdV-D48K (Fig. 9b).

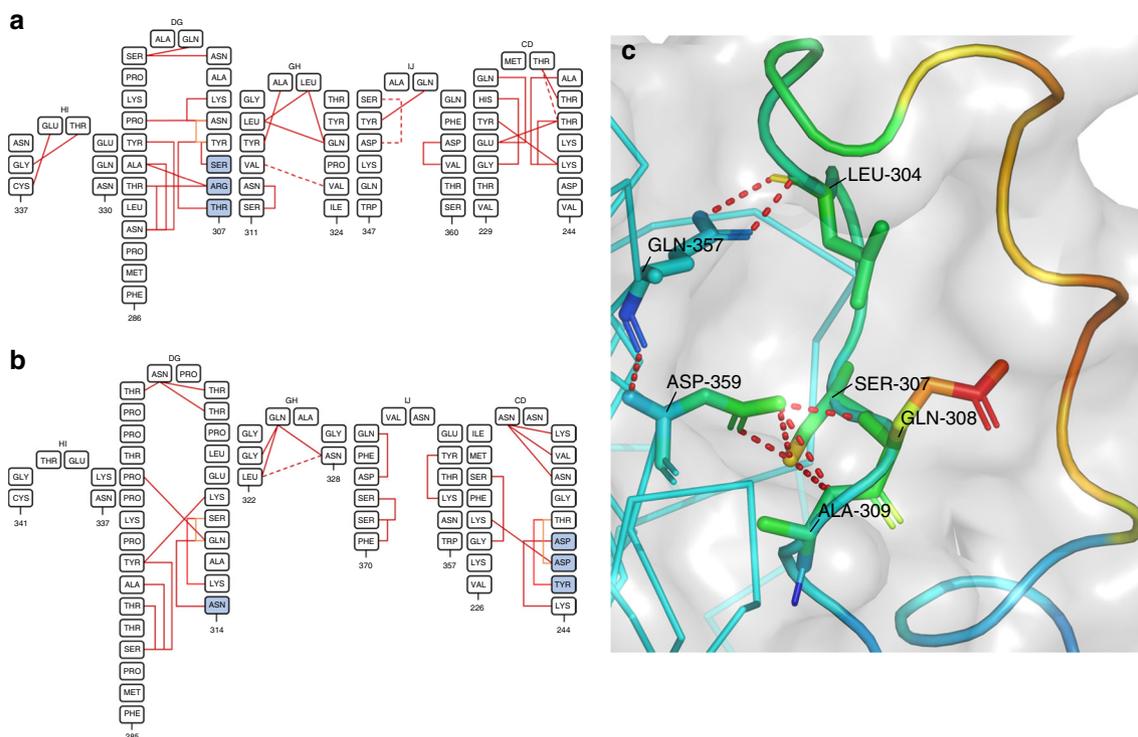


Fig. 4 Hypervariable loop conformations and contacts residues. The residues comprising the indicated loops of HAdV-D26K (**a**) and HAdV-D48K (**b**) are shown diagrammatically with numbers indicating the start and end residues of each loop depicted. The network of intraloop polar interactions is shown by solid lines (one polar bond), and dashed lines (two polar bonds, colour variations are only for ease of viewing) between interacting residues, similar interloop bonds are also present as visualised in Supplementary Figure 2. Residues forming part of a helical motif are shaded in blue. The HAdV-D48K DG-loop is seen to form contacts to the opposing monomer across the inter-monomer cleft (**c**). The labelled residues forming polar contacts (shown as sticks) are coloured by relative B-factor, with warmer colours indicating higher relative B-factors, as is the cartoon representation of the loop. The opposing HAdV-D48K monomer is seen as a ribbon representation of the carbon- α chain in cyan and the surface of the HAdV-D48K trimer seen as a semi-transparent grey surface

Discussion

This study reveals the crystal structure of two adenovirus proteins critical to primary receptor engagement, HAdV-D26 and HAdV-D48 fiber-knob, which are important viral vectors currently in human clinical trials^{3,11,12}. Despite their advanced development, the field lacks fundamental knowledge regarding the mechanisms of infection for these viral vector platforms. The work we described here provides a combined crystallographic, in silico, and in vitro approach to investigate adenovirus fiber-knob: receptor interactions with CAR and CD46, two receptors previously proposed to be utilised by these viruses^{5,9,17}.

Analysis of the phylogenetic relationship between 56 adenovirus serotypes, by both whole genome and fiber-knob domain alignment (Fig. 1a), confirms diversification into the widely accepted seven adenoviral species⁷. However, generating the phylogenetic tree with fiber-knob sequences, rather than whole genomes, shows additional diversity, not revealed by the whole virus taxonomy. Adenovirus species D breaks up into several additional sub-clades when focused on the fiber-knob, suggesting greater receptor diversity than might be expected based on the whole virus phylogeny. Similar observations have previously been made in species D hexon and penton¹³.

In contrast to species D, the phylogeny of species B adenoviruses, which are known to utilise Desmoglein 2 and CD46 as primary receptors, is simplified when focused upon the fiber-knob, indicating less diverse receptor usage^{39,43–45}. This simplification in comparison to the whole genomic alignment implies that much of the species diversity must lay in other proteins. The E3 protein, for example, is known to be highly diverse within

species B adenoviruses, having previously been exploited in the selection of the oncolytic (cancer killing) virus enadenotucirev, which is currently in clinical trials^{13,46,47}.

That we see such opposing effects on the species B and D phylogenetic trees when focusing on the fiber-knob, highlights the limitations of simple taxonomic approaches. The current adenoviral taxonomy is based on antibody neutralisation assays, which are limited by antibody's reliance on surface accessible proteins in the capsid, and does not account for diversity in other viral proteins, as the above suggests for species B. This supports a taxonomic proposal based upon viral genetics rather than antibody neutralisation, as has previously been suggested^{13,48}.

Many studies on adenovirus neutralisation have focused upon NAbs which bind to the hexon^{49–51}. Following intramuscular vaccination with non-replicating adenoviral vectors, most NAbs are targeted to the hexon; a reflection of its high abundance and surface availability in each viral capsid^{49,52}. However, during natural infection, many NAbs target the fiber protein⁵², presumably due to the abundance of fiber produced in the early stages of hAdV replication to loosen cell–cell junctions and facilitate viral spread, prior to lysis and entry of large amounts of whole virus to the bloodstream⁵³. For individuals with pre-existing anti-adenovirus immunity derived from natural infections, anti-fiber NAbs are likely to limit the use of vectors with common adenovirus fiber proteins by neutralisation of the viral vector prior to its therapeutic effect. Thus, for the development of vectors to circumvent pre-existing anti-adenovirus immunity for therapeutic use further exploration of this fiber protein diversity

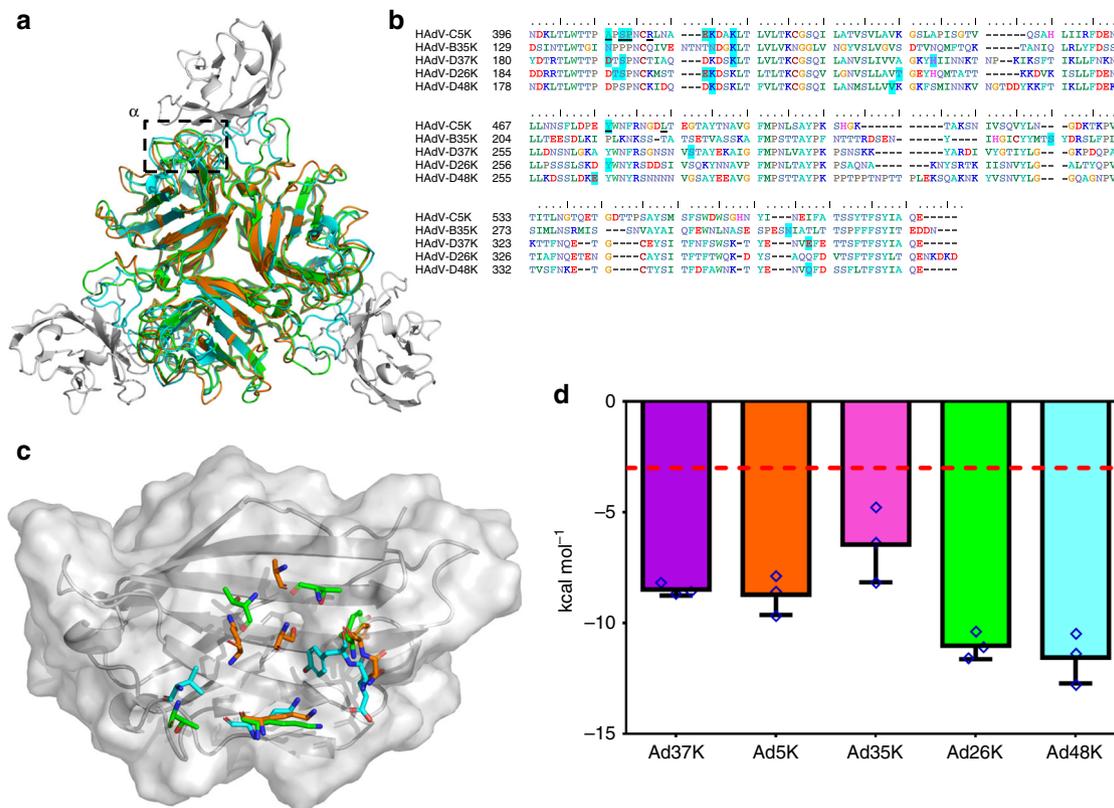


Fig. 5 Modelling of the HAdV-D26K and HAdV-D48K interaction with CAR at the α -interface. The α -interface region is shown by the box on the structural alignment of HAdV-C5K (orange), 26K (green), and 48K (cyan) fiber-knob domain crystal structures in complex with CAR-D1 domain (grey) as determined by homology alignment to the previously reported Ad37K CAR-D1 structure (PDB: 2J12) (**a**). The aligned amino acid sequence of the investigated fiber-knobs (**b**) and the predicted α -interface forming CAR-D1 binding residues are highlighted in blue, with the underlined residues representing the HAdV-C5K amino acids shown by Kirby et al. (2000)³⁷ to be important for CAR interaction. Conservation of key residues can be seen between HAdV-C5K, HAdV-D26K, and HAdV-D48K fiber-knobs. This conservation is visualised, with the contact residues comprising the α -interface with HAdV-C5K, 26K, and 48K shown as sticks in complex with the energy minimised CAR-D1 domain (grey), shown as the surface of the maximum spatial occupancy of the aligned CAR-D1 monomers from each of the energy minimised models in complex with HAdV-C5K, 26K, and 48K fiber-knobs (**c**). **d** Plots the predicted binding energy of the energy minimised fiber-knob proteins to CAR-D1 complex in the α -interface, only. Lower binding energy indicates a more stable interface with the red line depicting 3.0 kcal mol⁻¹, which can be considered background. $n = 3$, where each calculation is an independent fiber-knob: CAR interface, error bars indicate mean \pm SD

may be beneficial, as well as the on-going studies using hexon HVR pseudotypes to circumvent anti-hexon immunity^{11,54}.

Analysis of the adenovirus loops (Fig. 4, Supplementary Figure 3) reveals an intricate network of polar interactions which stabilise their three-dimensional structures. These bonds appear to hold the loops in a conformation which, in the case of HAdV-B11K and HAdV-B35K, facilitates receptor binding. In the HAdV-D26 and HAdV-D48 fiber-knob structures presented in this study the loops are also held in a stable conformation, though not one which enables CD46 interaction.

Modelling of HAdV-D26K/HAdV-D48K in complex with CD46 (Fig. 7) suggested few contacts, and interface energy calculations using these models predict a weak binding energy (Fig. 8). SPR indicated that HAdV-D26K has an affinity for CD46 that is approximately 1500 \times weaker than that of HAdV-B35K (Fig. 8d, e). Combined with the extremely fast kinetics, this is suggestive of a highly unstable interface. HAdV-D48K showed no affinity for CD46 at all. This was confirmed by in vitro competition inhibition assays, in which no tested quantity of recombinant fiber-knob was capable of inhibiting antibody binding to CD46 (Fig. 8c). These findings appear contradictory to previous studies which suggest CD46 as the primary receptor for these viruses^{16,17}. Our findings improve knowledge of the cell entry mechanisms of these viruses and the vectors derived from them,

and do not diminish the observed effectiveness of these vaccines. However, if CD46, a protein expressed on the surface of all nucleated cells, is not the receptor for these viruses then it is as yet unknown what the primary tissue tropism determinant is for these clinically significant viruses^{55,56}.

A similar methodology was applied to the interaction HAdV-D26K/HAdV-D48K with CAR. Inspection of the modelled complexes (Fig. 5) indicated a conserved α -interface enabling CAR binding in adenovirus 5, 26, and 48, fiber knobs. However, the structure of the β -interface interaction appears to indicate a mechanism modulating the fiber-knob's CAR affinity (Fig. 6). When occupying the intermonomer cleft in the conformation shown in Fig. 6b, the DG loops of HAdV-D26K and HAdV-D48K are likely to inhibit CAR binding by steric hindrance, but if the loops were to shift into a conformation which relieves this clash CAR binding could occur. Therefore, the ability of these vectors to interact with CAR is likely a function of the steric hindrance provided by these loops, reducing the ability of the fiber-knob domain to engage CAR in a permissive conformation. SPR analysis supports this hypothesis (Fig. 6d, e), where the larger the DG loop of the investigated fiber-knob the slower the K_{ON} .

The inter-loop contacts described in Fig. 4a, b, and Supplementary Figure 3, and the normalised B-factors described in Fig. 6b will influence the molecular dynamics of the DG-loops.

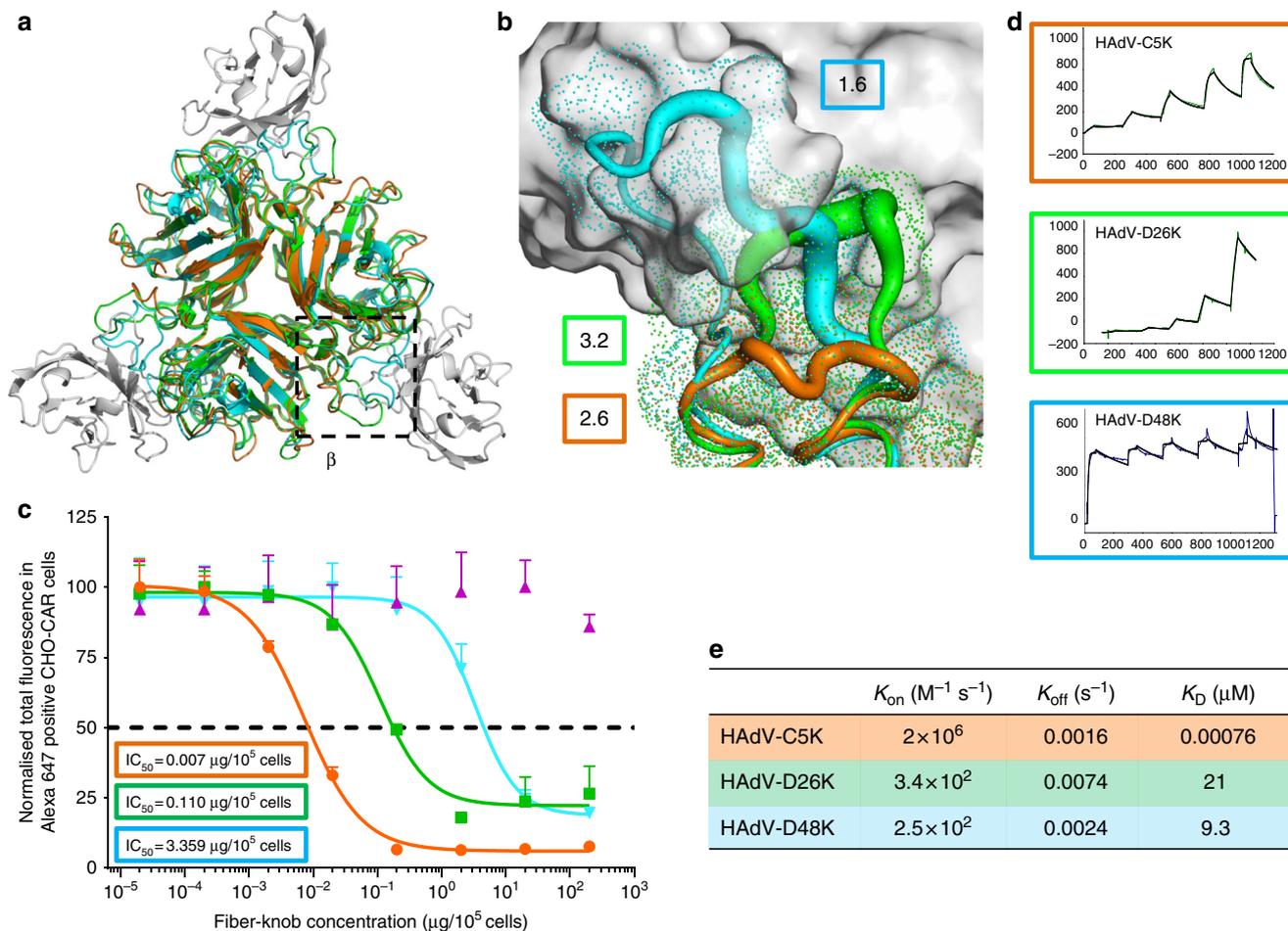


Fig. 6 Modelling of the HAAdV-D26K and HAAdV-D48K interaction with CAR at the β -interface. The β -interface region is shown by the box on the structural alignment of HAAdV-C5K (orange), 26K (green), and 48K (cyan) fiber-knob domain crystal structures in complex with CAR-D1 domain (grey) as determined by homology alignment to the previously reported Ad37K CAR-D1 structure (PDB: 2J12) (a). A dot surface shows the surface of HAAdV-C5K, 26K, and 48K DG-loops in the inter-monomer cleft (b). The boxes denote the maximum B-factor of the corresponding loops, which are shown as putty representations with thicker regions indicating higher relative B-factors, from which we can infer the relative stability of the loops. Antibody competition inhibition assay (c) shows the relative inhibitory ability of the HAAdV-C5, 35, 26, and 48, fiber-knob domains in CAR expressing CHO-CAR cells, with the calculated IC_{50} values shown in boxes. $n = 3$ biological replicates. Surface plasmon resonance (SPR) traces are shown by coloured lines, and the fitted curves by black lines (d). The calculated binding coefficients on rate (K_{On}), off rate (K_{Off}), and dissociation coefficient (K_D) are given in the table (e). IC_{50} curves are fitted by non-linear regression. Error bars represent standard deviation of 3 biological replicates. Error bars indicate mean \pm SD

Loops which can occupy a CAR inhibitory conformation but have fewer stabilising contacts, such as that of HAAdV-D26K (Supplementary Figure 3A, C), should be more permissive to CAR binding. While loops which are less flexible and/or stabilised in a CAR inhibitory conformation, such as HAAdV-D48K (Fig. 6b, Supplementary Figure 3B, D) should result in a fiber-knob which is less able to bind CAR. This hypothesis fits the competition inhibition studies shown in Fig. 6c, which demonstrate that HAAdV-D26K has an approximately $\sim 15\times$ lower affinity for CAR than HAAdV-C5K, and HAAdV-D48K has $500\times$ lower affinity.

Interestingly, the affinity of HAAdV-D48K for CAR as measured by SPR is approximately $2\times$ higher than that for HAAdV-D26K, due to the slower K_{Off} of HAAdV-D48K (Fig. 6e) which is in contrast with the IC_{50} curves (Fig. 6c) in which HAAdV-D48K is observed to bind to CAR less strongly than HAAdV-D26K. The incongruity may be explained by the methodology. It is possible that the large fluid volume in the wells during the inhibition assay (in comparison to the BIAcore microfluidics system), favoured greater binding by HAAdV-D26K due to its faster K_{On} , compared to HAAdV-D48K. This discrepancy does not alter the proposed

model of CAR interaction, and seems to confirm the importance of the K_{On} , presumably mediated by the β -interface.

Species D adenoviruses have a large range of different DG loops (Supplementary Figure 5). Most sequences have lengths equal to, or greater than, that of HAAdV-D26K, making it plausible that they too could modulate the fiber-knob interaction with CAR. However, the magnitude of this effect will be dependent on the individual molecular dynamics of the DG-loops and its interactions with adjacent residues.

Assuming this mechanism of CAR binding regulation is broadly applicable, it may have important implications for adenoviral vector design. The presence of a high affinity receptor for the virus can mask the low affinity CAR interaction, creating a hidden tropism only observed if the virus is forced to rely upon it. Expression of CAR on human erythrocytes suggests the potential for sequestration of virotherapies in the blood⁵⁷. CAR expression in lung epithelial tissues offers another site for potential off target activity^{58,59}. Therefore, many virotherapies previously thought to be non-CAR binding adenoviruses may in fact demonstrate weak CAR tropism, driving off target infections or resulting in

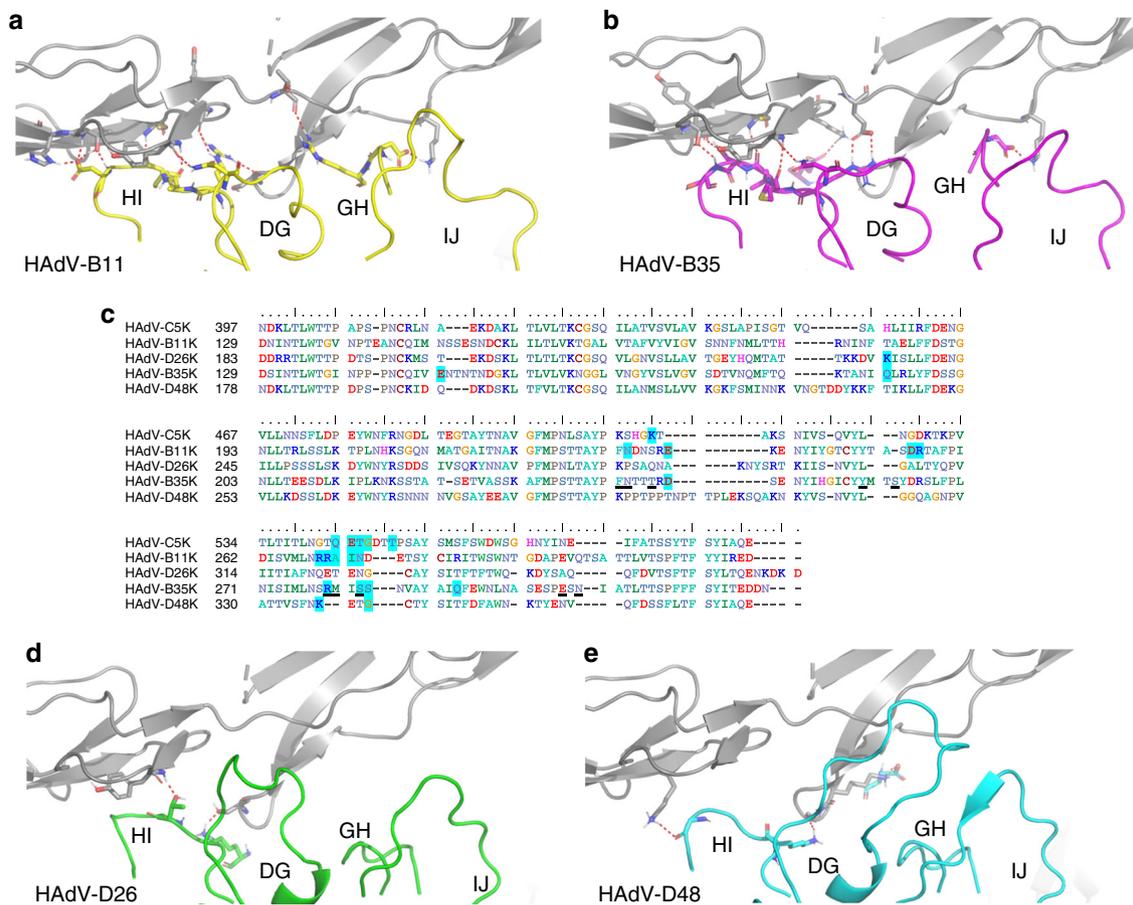


Fig. 7 Modelling of the HAAdV-D26K and HAAdV-D48K with CD46. Red dashes show contacts between the energy minimised crystal structure of CD46 SC1 and SC2 domains (grey cartoon) and Ad11K in complex (PDB 3O8E) (**a**). The known CD46 interacting fiber-knob, HAAdV-B35K (purple), is aligned to the above crystal structure and energy minimised (**b**). Amino acid sequence alignment of the tested fiber-knob proteins (**c**) shows conservation of residues previously shown by Wang et al. (2007)³⁶ to be key to CD46 binding (underlined) between the known CD46 binding fiber-knobs, Ad11K and HAAdV-B35K. Residues highlighted in blue are predicted to form direct contacts with CD46 in the energy minimised models. Similar alignments to that performed with HAAdV-B35K are shown for HAAdV-D26K (Green—**d**) and HAAdV-D48K (cyan—**e**). In all models, red dashes indicate polar contacts between the residues shown as stick representations

sequestration of the vector in tissues other than that target. This may not be of grave consequence for non-replicating vectors, such as viral vaccines, but in vectors which rely upon controlled replication in targeted tissues, such as oncolytic virotherapies, this could result in off-target infection, dysregulated expression of therapeutic protein, and reduced delivery to the point of need.

DSG2 was also shown to be unable to bind HAAdV-D26K or HAAdV-D48K at any concentration by SPR. It is notable that the K_D measured for the HAAdV-B3K (66.9 μ M) is much lower than that measured during the original investigation of DSG2 as an HAAdV-B3K receptor (2.3 nM)⁴⁴. This is likely due to our use of recombinant knob trimers, rather than the multivalent penton dodecahedrons.

The final, known, adenovirus fiber-knob receptor, which has thus far not been addressed in this study is sialic acid, as part of glycosylation motifs. Several adenoviruses have been shown to bind to sialylated glycans, including HAAdV-D37^{32,38}, HAAdV-G52^{30,60}, and Canine adenovirus serotype 2 (CAAdV-2)⁵⁷. Each of these three viruses binds to sialic acid by different mechanisms (Supplementary Figure 6). Supplementary Figure 6 shows that HAAdV-D26/48K do not share the sialic acid binding residues found in HAAdV-G52K or CAAdV-2 but do share the Tyr-142 and Lys-178 contact residues with HAAdV-D37K. Further, the HAAdV-D37K contact residue Pro-147 is between the sialic acid and the main chain oxygen which is functionally identical at the similar position in HAAdV-D26/48K. Taken

together, it remains plausible that HAAdV-D26/48K may be capable of binding sialic acid in an HAAdV-D37K-like manner. However, binding does not equate to functional infection, as seen with HAAdV-D19pK³² and further studies are required to ascertain whether HAAdV-D26/48 are capable of utilising sialic acid to generate a productive infection. Further, HAAdV-D37 was shown to require a specific glycosylation motif (GD1a) in order form a functional infection, so any assessment of sialic acid as an adeno-viral receptor must be in the context of its glycan carrier³⁸.

The work undertaken in this study presents, for the first time, the crystal structures of the fiber-knob domain protein of HAAdV-D26 (PDB: 6FJN), and HAAdV-D48 (PDB: 6FJQ) fiber-knob protein. In addition, we report a new crystal structure for HAAdV-C5 fiber-knob protein (PDB: 6HCN) with improved resolution compared to the existing structure (PDB: 1KNB)²⁷. We utilised these structures to investigate the ability of these proteins to interact with the putative receptors, CAR and CD46, by an integrative structural, in silico, and in vitro workflow. We demonstrate that HAAdV-D26 and HAAdV-D48 fiber-knob domains have a weak ability to bind CAR, and negligible CD46 interaction, suggesting that these viruses are unlikely to utilise these proteins as a primary receptor in vivo. Finally, we showed that DSG2 is also unable to form a stable interaction in the context of SPR analysis. We suggest that CAR binding is mod-erated by a previously unreported mechanism of steric inhibition

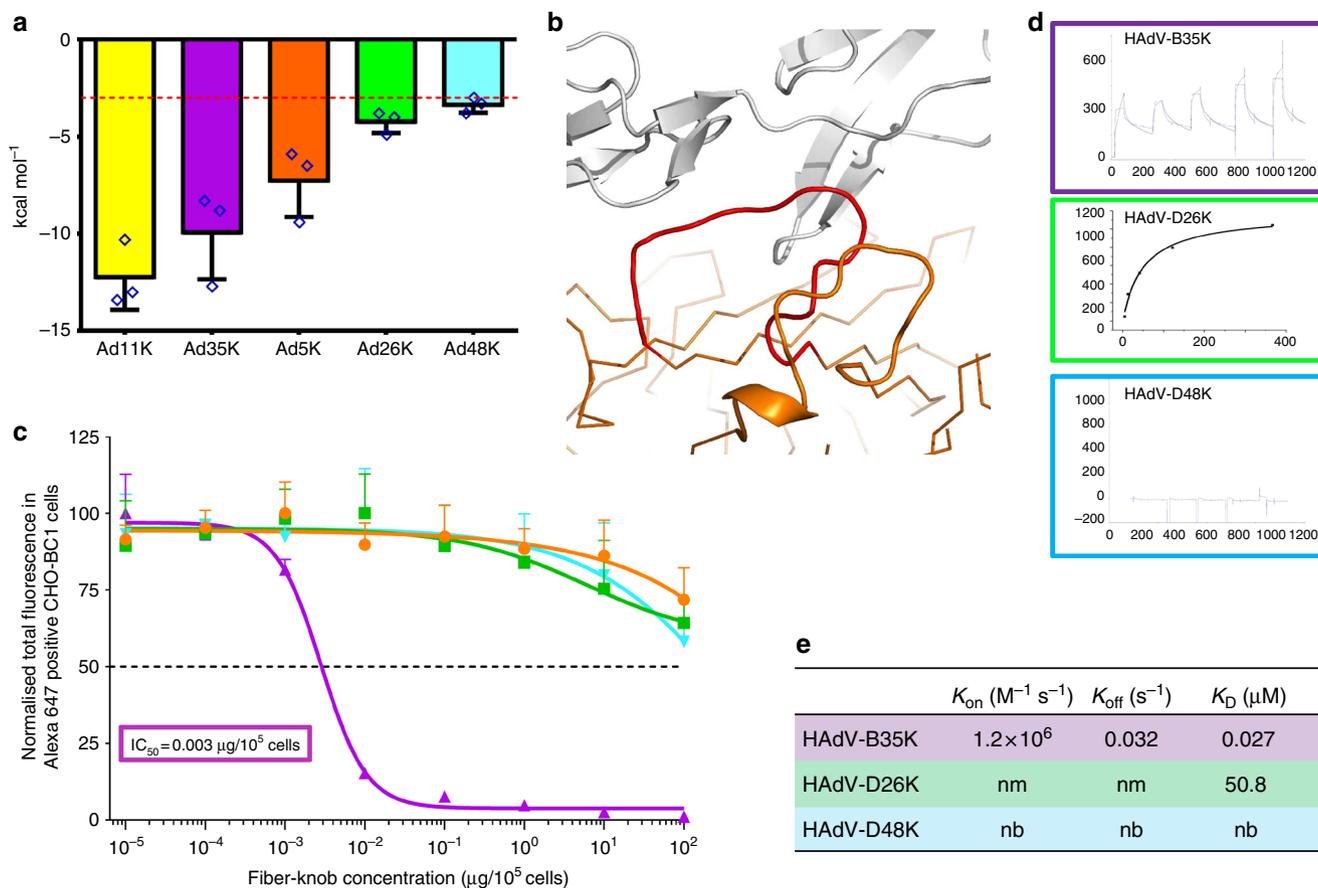


Fig. 8 Binding energetics and affinities of HAdV-D26K and HAdV-D48K with CD46. Calculation of the predicted binding energies for the energy minimised fiber-knob: CD46 models are compared on the bar chart (a), lower kcal mol⁻¹ values indicate a stronger interaction, the red line at 3.0 kcal mol⁻¹, denotes an interface energy which can be considered negligible (random proteins passing in solution), $n = 3$, where each calculation is an independent fiber-knob: CD46 interface. The HI loop (red) of the HAdV-C5 fiber-knob (orange) is seen to extend between CD46 (grey) and the DG loop (b). The antibody competition inhibition assay (c) shows the relative inhibitory ability of the HAdV-C5, HAdV-B35, HAdV-D26, and HAdV-D48 fiber-knob domains in CD46 expressing CHO-BC1 cells, with the calculated IC₅₀ values shown in boxes. $n = 3$ biological replicates. Surface plasmon resonance (SPR) traces are shown by coloured lines, and the fitted curves by black lines (d). The calculated binding coefficients on rate (K_{on}), off rate (K_{off}), and dissociation coefficient (K_D) are given in the table (e), nm (not measured) indicates that the kinetics were too fast to measure, nb denotes no binding. IC₅₀ curves are fitted by non-linear regression. Error bars represent standard deviation of 3 biological replicates. Error bars indicate mean ± SD

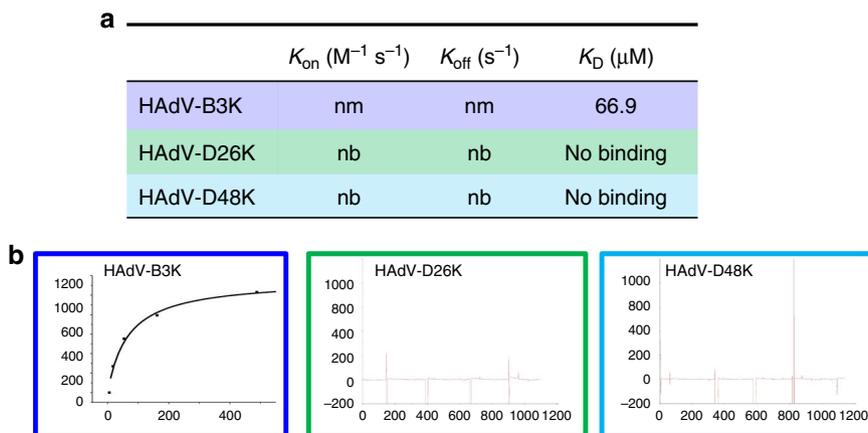


Fig. 9 Desmoglein 2 is unlikely to be a receptor for HAdV-D26K or HAdV-D48K. The dissociation constant was calculated for HAdV-B3K binding to DSG2, but kinetics were too fast to determine K_{on} or K_{off} (a), the K_D curve is shown for HAdV-B3K while HAdV-D26K and HAdV-D48K are seen to form no interaction with DSG2 (b). nm (not measured) indicates that the kinetics were too fast to measure, nb denotes no binding

which may apply to other adenoviruses and demonstrate an *in silico* methodology capable of rapidly predicting these interactions. These findings enhance our understanding of the virology of adenovirus infection, and have direct implications for virotherapy vector design, which often rely upon carefully controlled receptor tropisms to achieve specificity and efficacy^{9,18,28}.

Methods

Genome alignment and analysis of genetic diversity. Representative whole genomes (nucleotide) of adenoviral species 1–56 were selected from the National Center for Biotechnology Information (NCBI), and aligned using the EMBL-EBI Clustal Omega tool^{61,62}. Fiber-knob domain amino acid sequences were derived from the same genome sequences, defined as the translated nucleotide sequence of the fiber protein (pIV) from the conserved TLW hinge motif to the protein C-terminus, and aligned in the same manner as the whole genomes. Phylogenetic relationships were inferred using the maximum likelihood method based upon the Jukes Cantor model for the whole genome nucleotide analysis⁶³, and the Poisson correction model for the fiber-knob amino acid analysis⁶⁴, using MEGA X software⁶⁵. Confidence was determined by bootstrap analysis (500 replicates)⁶⁶ and trees displayed condensed at 70% confidence (percentage confidence values shown at each node) where stated.

Fiber-knob amino acid variability. Amino acid sequence variability scores were calculated from the Clustal Omega aligned amino acid sequences of the fiber-knob domains of either adenoviruses 1–56, or only the species D adenoviruses. Analysis was performed using the Protein Variability Server (PVS), using a consensus base sequence and the Wu–Kabat method⁶⁷.

Generation of recombinant fiber-knob protein. SG13009 *Escherichia coli* harbouring pREP-4 plasmid and pQE-30 expression vector containing the fiber-knob DNA sequence were cultured in 20 ml LB broth with 100 µg/ml ampicillin and 50 µg/ml kanamycin overnight from glycerol stocks made in previous studies^{18,68,69}. 1 L of TB (Terrific Broth, modified, Sigma-Aldrich) containing 100 µg/ml ampicillin and 50 µg/ml were inoculated with the overnight *E. coli* culture and incubated at 37 °C until they reached OD_{0.6}. IPTG was then added to a final concentration of 0.5 mM and the culture incubated at 37 °C for 4 h. Cells were then harvested by centrifugation at 3000g, resuspended in lysis buffer (50 mM Tris, pH 8.0, 300 mM NaCl, 1% (v/v) NP40, 1 mg/ml Lysozyme, 1 mM β-mercaptoethanol), and incubated at room temperature for 30 min. Lysate was clarified by centrifugation at 30,000g for 30 min and filtered through a 0.22 µm syringe filter (Millipore, Abingdon, UK). Clarified lysate was then loaded onto a 5 ml HisTrap FF nickel affinity chromatography column (GE) at 2.0 ml/min and washed with 5 column volumes into elution buffer A (50 mM Tris [pH 8.0], 300 mM NaCl, 1 mM β-mercaptoethanol). Protein was eluted by 30 min gradient elution from buffer A to B (buffer A + 400 mM 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.

Competition inhibition assays. CHO cells expressing the appropriate receptor (CAR: CHO-CAR, or CD46: CHO-BC1) were seeded at a density of 30,000 cells per well in a flat bottomed 96-well tissue culture plate and incubated at 37 °C overnight. Serial dilutions were made up in serum-free RPMI-1640 to give a final concentration range of 0.0001–100 µg/10⁵ cells of recombinant soluble knob protein. Cells were incubated on ice for 30 min, then washed twice with cold PBS. Fiber-knob dilutions were then added to the cells and incubated on ice for 30 min. Cells were then washed twice in cold PBS and stained with the primary CAR or CD46 antibody, RmcB (Millipore; 05-644) or MEM-258 (Abcam; Ab789), respectively, to complex receptors unbound by fiber-knob protein, and incubated for 1 h on ice. Cells were washed twice further in PBS and incubated on ice for 1 h with Alexa-647 labelled goat anti-mouse F(ab')₂ (ThermoFisher; A-21237)^{18,68,69}. All antibodies were used at a concentration of 2 µg/ml.

Samples were run in triplicate and analysed by flow cytometry on Attune NxT (ThermoFisher), and analysed using FlowJo v10 (FlowJo, LLC) by gating sequentially on singlets, cell population, and Alexa-647 positive cells. Total fluorescence (TF) was used as the measure of inhibition, rather than percentage of fluorescent cells in the total population, to account for the presence of multiple receptor copies per cell surface which can enable partial inhibition of antibody binding on the cell surface. TF was defined as the percentage of Alexa-647 positive cells in the single cell population for each sample and multiplied by the median fluorescent intensity (MFI) of the Alexa-647 positive single cell population in each sample. Data points are the mean total fluorescence of three biological replicates with error given as the standard deviation from the mean. IC₅₀ curves were fitted by non-linear regression, and used to determine the IC₅₀ concentrations^{18,68,69}. CHO-CAR and CHO-BC1 cells were originally derived by Bergelson et al.⁷⁰ and Manchester et al.⁷¹, respectively.

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 7.5 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 200 nl of screen solution and 200 nl of protein solution were equilibrated against a reservoir of 60 µl crystallisation solution. The plates were sealed and incubated at 18 °C.

Crystals of HAdV-C5K appeared in PACT Premier condition D04 (0.1 M MMT, pH 7.0, 20% PEG 1500), within 1–7 days. Crystals of HAdV-D26K appeared within 1–7 days, in PACT Premier (Molecular Dimensions, Suffolk, UK) condition A04; 0.1 M MMT [DL-malic acid, MES monohydrate, Tris], pH 6.0, 25% PEG 1500. Crystals of HAdV-D48K appeared in PACT Premier condition D02 (0.1 M Bis-Tris-propane, pH 6.5, 20% PEG 3350, 0.2 M NaNO₃), within 2 weeks. 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 0.9795 Å. 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, Xia⁷², DIALS, and Autoproc⁷³. Scaling and merging data was completed with Pointless, Aimless and Truncate from the CCP4 package⁷⁴. Structures were solved with PHASER⁷⁵, COOT⁷⁶ was used to correct the sequences and adjust the models, REFMAC5⁷⁷ was used to refine the structures and calculate maps. Graphical representations were prepared with PyMOL⁷⁸. Reflection data and final models were deposited in the PDB database with accession codes: HAdV-C5K, 6HCN; HAdV-D26k, 6FJN; and HAdV-D48k, 6FJQ. Full crystallographic refinement statistics are given in Supplementary Table 2; stereo images depicting representative areas of the model and map are provided in Supplementary Figure 7.

Modelling of fiber-knob ligand interactions. Fiber-knob proteins were modelled in complex with CAR or CD46 using the existing HAdV-D37K—CAR liganded (PDB 2J12) or the HAdV-B11K—CD46 liganded (PDB 3O8E) structures, respectively, as a template. Non-protein components and hydrogens were removed from the template model and the fiber-knob protein of interest. The two fiber-knob proteins were then aligned with respect to their Ca chains, in such a way as to achieve the lowest possible RMSD. Models containing only the fiber-knob protein of interest and the ligand were saved and subjected to energy minimisation, using the YASARA self-parametrising energy minimisation algorithm as performed by the YASARA energy minimisation server, and results were visualised in PyMol^{78,79}.

Calculation of interface energy. Interface energies were calculated using QT-PISA using biological protein assemblies and excluding crystallographic interfaces⁸⁰. Values are the mean of the three symmetrical interfaces in each trimer and error is the standard deviation from the mean, any values above -3.0 kcal mol⁻¹ were considered to be background as shown as a red dashed line on graphs⁴².

Sequence alignments. Alignments were performed using the Clustal Omega multiple sequence alignment algorithm and visualised with BioEdit^{61,62}.

B-factor normalisation. Comparing order between different structures by comparing individual B-factors can be misleading. Post-refinement B-factors relate to the Wilson B-factor, which can vary widely between data sets, even from the same crystal preparation. A valid comparison between different structures can be achieved by comparing normalised B-factors instead. Normalisation was performed by dividing individual atomic B-factors by the average B-factor of the whole data set, quantifying the range of internal flexibility in a structure. This normalised B-factor can then be compared between different data sets.

Surface plasmon resonance (SPR) analysis. Binding analysis was performed using a BIACore 3000™ equipped with a CM5 sensor chip. Approximately 5000 RU of CD46, CAR, and DSG2 was attached to the CM5 sensor chip, using amine coupling, at a slow flow-rate of 10 µl/min to ensure uniform distribution on the chip surface. A blank flow cell was used as a negative control surface on flow cell 1. All measurements were performed at 25 °C in PBS buffer (Sigma, UK) at a flow rate of 30 µl/min. For equilibrium binding analysis, the HAdV-D26K and HAdV-B3K fiber knob proteins were purified and concentrated to 367 and 3 µM respectively. 5 × 1:3 serial dilutions were prepared for each sample and injected over the relevant sensor chip. The equilibrium binding constant (K_D) values were calculated assuming a 1:1 interaction by plotting specific equilibrium-binding responses against protein concentrations followed by non-linear least squares fitting of the Langmuir binding equation. For single cycle kinetic analysis, HAdV-D26K, HAdV-D48K, HAdV-B35K, HAdV-C5K, and HAdV-B3K were injected at a top concentration of around 200 µM, followed by four injections using serial 1:3 dilutions. The K_D values were calculated assuming Langmuir binding (AB = B × ABmax / (K_D + B)) and the data were analysed using the kinetic titration algorithm (BIAevaluationTM 3.1). Receptor proteins were obtained commercially, as follows: Recombinant Human Desmoglein-2 Fc Chimera Protein, R&D Systems, Catalogue number 947-DM-100. Recombinant Human CXADR Fc Chimera Protein (CAR),

R&D Systems, Catalogue number 3336-CX-050. Recombinant Human CD46 Protein (His Tag), Sino Biological, Catalogue number 12239-H08H.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Macromolecular structures generated during this study have been deposited in wwPDB (worldwide Protein Data Bank; <https://www.wwpdb.org/>), and have PDB ID's 6FJN, 6HCN, and 6FJQ. PDB ID's for macromolecular structures utilised, but not generated in the course of this study, are as follows: HAdV-B11K in complex with CD46, PDB 3O8E. HAdV-D37K in complex with CAR-D1, PDB 2J12. HAdV-B35K, PDB 2QLK. Genomic sequences from which fiber-knob domain sequences were determined, which have been used in phylogenetic analysis, have the following NCBI accession numbers: AC_000017|AF532578|X73487|AY803294|AB562586|AY601636|AF108105|GU191019|JQ326209|AC_000007|JN226749|KF528688|FJ404771|JN226750|JN226751|JN226752|EF153474|JN226753|FJ824826|JN226754|JN226755|AM749299|JN226756|JN226758|AY737797|AC_000019|GQ384080|JN226759|JN226760|KU162869|DQ315364|JN226761|JN226762|JN226763|JN226764|AY875648|JN226757|EF153473|DQ393829|AC_000008|AY737798|JN226765|DQ923122|AB605243|NC_012959|FJ643676|HM770721|HQ413315|AC_000018|DQ086466|JN226746|JN226747|AB448776|AB448767|AJ854486|KF006344|. All other data pertaining to this manuscript are available from the authors upon request.

Received: 22 August 2018 Accepted: 21 January 2019

Published online: 14 February 2019

References

1. Combination Adenovirus+Pembrolizumab to Trigger Immune Virus Effects—Full Text View—ClinicalTrials.gov. Available at: <https://clinicaltrials.gov/ct2/show/NCT02798406?term=adenovirus&recrs=a&cond=Glioblastoma&rank=2> (accessed 23rd September 2017).
2. A Safety and Efficacy Study of ChAdOx1 LS2 and MVA LS2—Full Text View—ClinicalTrials.gov. Available at: <https://clinicaltrials.gov/ct2/show/NCT03203421?term=adenovirus&recrs=a&cond=malaria&rank=2> (accessed 23rd September 2017).
3. 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> (accessed 23rd September 2017).
4. Lee, C. S. et al. Adenovirus-mediated gene delivery: potential applications for gene and cell-based therapies in the new era of personalized medicine. *Genes Dis.* **4**, 43–63 (2017).
5. Abbink, P. 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).
6. King, A. M. Q., Adams, M. J., Carstens, E. B. & Lefkowitz, E. J. (eds). Family—Adenoviridae. in *Virus Taxonomy* 125–141 (Elsevier, 2012). <https://doi.org/10.1016/B978-0-12-384684-6.00009-4>.
7. 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, 1995).
8. Mast, T. C. et al. International epidemiology of human pre-existing adenovirus (Ad) type-5, type-6, type-26 and type-36 neutralizing antibodies: correlates of high Ad5 titers and implications for potential HIV vaccine trials. *Vaccine* **28**, 950–957 (2010).
9. Uusi-Kerttula, H., Hulin-Curtis, S., Davies, J. & Parker, A. L. Oncolytic adenovirus: strategies and insights for vector design and immuno-oncolytic applications. *Viruses* **7**, 6009–6042 (2015).
10. Priddy, F. H. et al. Safety and immunogenicity of a replication-incompetent adenovirus type 5 HIV-1 Clade B *gag/pol/nef* vaccine in healthy adults. *Clin. Infect. Dis.* **46**, 1769–1781 (2008).
11. Baden, L. R. 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).
12. Barouch, D. H. 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). *Lancet* **392**, 232–243 (2018).
13. Robinson, C. M. et al. Molecular evolution of human adenoviruses. *Sci. Rep.* **3**, 1812 (2013).
14. Rux, J. J., Kuser, P. R. & Burnett, R. M. Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution X-ray crystallographic, molecular modeling, and sequence-based methods. *J. Virol.* **77**, 9553–9566 (2003).
15. Singh, G. et al. Overreliance on the hexon gene, leading to misclassification of human adenoviruses. *J. Virol.* **86**, 4693–4695 (2012).
16. Teigler, J. E., Iampietro, M. J. & Barouch, D. H. Vaccination with adenovirus serotypes 35, 26, and 48 elicits higher levels of innate cytokine responses than adenovirus serotype 5 in Rhesus monkeys. *J. Virol.* **86**, 9590–9598 (2012).
17. Li, H. et al. Adenovirus serotype 26 utilizes CD46 as a primary cellular receptor and only transiently activates T lymphocytes following vaccination of rhesus monkeys. *J. Virol.* **86**, 10862–10865 (2012).
18. Uusi-Kerttula, H. et al. Pseudotyped $\alpha v\beta 6$ integrin-targeted adenovirus vectors for ovarian cancer therapies. *Oncotarget* **7**, 27926–27937 (2016).
19. Camacho, Z. T., Turner, M. A., Barry, M. A. & Weaver, E. A. CD46-mediated transduction of a species D adenovirus vaccine improves mucosal vaccine efficacy. *Hum. Gene Ther.* **25**, 364–374 (2014).
20. Weaver, E. A. & Barry, M. A. Low seroprevalent species D adenovirus vectors as influenza vaccines. *PLoS ONE* **8**, e73313 (2013).
21. Geisbert, T. W. 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).
22. Wang, Y. et al. Phylogenetic evidence for intratypic recombinant events in a novel human adenovirus C that causes severe acute respiratory infection in children. *Sci. Rep.* **6**, 23014 (2016).
23. Ebner, K., Pinsker, W. & Lion, T. Comparative sequence analysis of the hexon gene in the entire spectrum of human adenovirus serotypes: phylogenetic, taxonomic, and clinical implications. *J. Virol.* **79**, 12635–12642 (2005).
24. Al Qurashi, Y. M. A., Alkhalaf, M. A., Lim, L., Guiver, M. & Cooper, R. J. Sequencing and phylogenetic analysis of the hexon, fiber, and penton regions of adenoviruses isolated from AIDS patients. *J. Med. Virol.* **84**, 1157–1165 (2012).
25. Pache, L., Venkataraman, S., Nemerow, G. R. & Reddy, V. S. Conservation of fiber structure and CD46 usage by subgroup B2 adenoviruses. *Virology* **375**, 573–579 (2008).
26. Russell, W. C. Adenoviruses: update on structure and function. *J. Gen. Virol.* **90**, 1–20 (2009).
27. Xia, D., Henry, L. J., Gerard, R. D. & Deisenhofer, J. Crystal structure of the receptor-binding domain of adenovirus type 5 fiberprotein at 1.7 Å resolution. *Structure* **2**, 1259–1270 (1994).
28. Baker, A. T., Aguirre-Hernández, C., Halldén, G. & Parker, A. L. Designer oncolytic adenovirus: coming of age. *Cancers* **10**, 201 (2018).
29. Amoureux, M.-C. 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).
30. Lenman, A. 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).
31. Arnberg, N. et al. Adenovirus type 37 binds to cell surface sialic acid through a charge-dependent interaction. *Virology* **302**, 33–43 (2002).
32. Arnberg, N., Edlund, K., Kidd, A. H. & Wadell, G. Adenovirus type 37 uses sialic acid as a cellular receptor. *J. Virol.* **74**, 42–48 (2000).
33. Cupelli, K. et al. Structure of adenovirus type 21 knob in complex with CD46 reveals key differences in receptor contacts among species B adenoviruses. *J. Virol.* **84**, 3189–3200 (2010).
34. Persson, B. D. et al. Structure of the extracellular portion of CD46 provides insights into its interactions with complement proteins and pathogens. *PLoS Pathog.* **6**, e1001122 (2010).
35. Persson, B. D. et al. Adenovirus type 11 binding alters the conformation of its receptor CD46. *Nat. Struct. Mol. Biol.* **14**, 164–166 (2007).
36. Wang, H. et al. Identification of CD46 binding sites within the adenovirus serotype 35 fiber knob. *J. Virol.* **81**, 12785–12792 (2007).
37. Kirby, I. et al. Identification of contact residues and definition of the CAR-binding site of adenovirus type 5 fiber protein. *J. Virol.* **74**, 2804–2813 (2000).
38. Nilsson, E. C. et al. The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. *Nat. Med.* **17**, 105–109 (2011).
39. Fleischli, C. 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).
40. Seiradake, E., Lortat-Jacob, H., Billet, O., Kremer, E. J. & Cusack, S. 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).
41. Krieger, E., Koraimann, G. & Vriend, G. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins Struct. Funct. Bioinformatics* **47**, 393–402 (2002).
42. Jensen, J. H. Predicting accurate absolute binding energies in aqueous solution: thermodynamic considerations for electronic structure methods. *Phys. Chem. Chem. Phys.* **17**, 12441–12451 (2015).
43. Vassal-Stermann, E. 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**, 8381 (2018).
44. Wang, H. et al. Desmoglein 2 is a receptor for adenovirus serotypes 3, 7, 11 and 14. *Nat. Med.* **17**, 96–104 (2011).

45. Segerman, A., Arnberg, N., Erikson, A., Lindman, K. & Wadell, G. There are two different species B adenovirus receptors: sBAR, common to species B1 and B2 adenoviruses, and sB2AR, exclusively used by species B2 adenoviruses. *J. Virol.* **77**, 1157–1162 (2003).
46. Kuhn, I. et al. Directed evolution generates a novel oncolytic virus for the treatment of colon cancer. *PLoS ONE* **3**, e2409 (2008).
47. Illingworth, S. et al. Preclinical safety studies of enadenotucirev, a chimeric group B human-specific oncolytic adenovirus. *Mol. Ther. Oncolytics* **5**, 62–74 (2017).
48. HAdV Working Group. <http://hadv.wg.gmu.edu/>.
49. Bradley, R. R., Lynch, D. M., Iampietro, M. J., Borducchi, E. N. & Barouch, D. H. Adenovirus serotype 5 neutralizing antibodies target both hexon and fiber following vaccination and natural infection. *J. Virol.* **86**, 625–629 (2012).
50. Bradley, R. R. et al. Adenovirus serotype 5-specific neutralizing antibodies target multiple hexon hypervariable regions. *J. Virol.* **86**, 1267–1272 (2012).
51. Ma, J. et al. Manipulating adenovirus hexon hypervariable loops dictates immune neutralisation and coagulation factor X-dependent cell interaction in vitro and in vivo. *PLoS Pathog.* **11**, e1004673 (2015).
52. Parker, A. L. et al. Effect of neutralizing sera on factor x-mediated adenovirus serotype 5 gene transfer. *J. Virol.* **83**, 479–483 (2009).
53. Walters, R. W. et al. Adenovirus fiber disrupts CAR-mediated intercellular adhesion allowing virus escape. *Cell* **110**, 789–799 (2002).
54. Keele, B. F. et al. Adenovirus prime, Env protein boost vaccine protects against neutralization-resistant SIVsmE660 variants in rhesus monkeys. *Nat. Commun.* **8**, 15740 (2017).
55. Andrews, P. W. et al. A human cell-surface antigen defined by a monoclonal antibody and controlled by a gene on human chromosome 1. *Ann. Hum. Genet.* **49**, 31–39 (1985).
56. Uhlen, M. et al. A pathology atlas of the human cancer transcriptome. *Science* **357**, eaan2507 (2017).
57. Seiradake, E. et al. The cell adhesion molecule “CAR” and sialic acid on human erythrocytes influence adenovirus in vivo biodistribution. *PLoS Pathog.* **5**, e1000277 (2009).
58. Cohen, C. J. et al. The coxsackievirus and adenovirus receptor is a transmembrane component of the tight junction. *Proc. Natl. Acad. Sci. USA* **98**, 15191–15196 (2001).
59. Awasthi, V., Meinken, G., Springer, K., Srivastava, S. C. & Freimuth, P. Biodistribution of radiolabeled adenovirus fiber protein knob domain after intravenous injection in mice. *J. Virol.* **78**, 6431–6438 (2004).
60. Lenman, A. et al. Polysialic acid is a cellular receptor for human adenovirus 52. *Proc. Natl. Acad. Sci. USA* <https://doi.org/10.1073/pnas.1716900115> (2018).
61. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539–539 (2011).
62. Li, W. et al. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* **43**, W580–W584 (2015).
63. Jukes, T. H. & Cantor, C. R. Evolution of protein molecules. in *Mammalian Protein Metabolism* 21–132 (Elsevier, 1969). <https://doi.org/10.1016/B978-1-4832-3211-9.50009-7>.
64. Zuckerkandl, E. & Pauling, L. Evolutionary divergence and convergence in proteins. in *Evolving Genes and Proteins* (eds Bryson, V. & Vogel, H. J.) 97–166 (Academic Press, 1965). <https://doi.org/10.1016/B978-1-4832-2734-4.50017-6>.
65. Kumar, S., Stecher, G., Li, M., Niyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
66. Felsenstein, J. Confidence limits on phylogenies: an approach using the Bootstrap. *Evolution* **39**, 783–791 (1985).
67. Garcia-Boronat, M., Diez-Rivero, C. M., Reinherz, E. L. & Reche, P. A. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Res.* **36**, W35–W41 (2008).
68. Coughlan, L. et al. Retargeting adenovirus serotype 48 fiber knob domain by peptide incorporation. *Hum. Gene Ther.* **25**, 385–394 (2014).
69. Coughlan, L. et al. In vivo retargeting of adenovirus type 5 to $\alpha v\beta 6$ integrin results in reduced hepatotoxicity and improved tumor uptake following systemic delivery. *J. Virol.* **83**, 6416–6428 (2009).
70. Bergelson, J. M. et al. Isolation of a common receptor for coxsackie B viruses and adenoviruses 2 and 5. *Science* **275**, 1320–1323 (1997).
71. Manchester, M., Liszewski, M. K., Atkinson, J. P. & Oldstone, M. B. Multiple isoforms of CD46 (membrane cofactor protein) serve as receptors for measles virus. *Proc. Natl. Acad. Sci. USA* **91**, 2161–2165 (1994).
72. Winter, G., Lobley, C. M. C. & Prince, S. M. Decision making in xia2. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1260–1273 (2013).
73. Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 293–302 (2011).
74. Dodson, E. J., Winn, M. & Ralph, A. Collaborative computational project, number 4: providing programs for protein crystallography, Vol. 277. in *Methods in Enzymology* 620–633 (Academic Press, 1997).
75. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
76. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
77. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
78. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 2.0.
79. Krieger, E. et al. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins Struct. Funct. Bioinformatics* **77**, 114–122 (2009).
80. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

Acknowledgements

A.T.B. is supported by a Tenovus Cancer Care PhD studentship to A.L.P. (reference PhD2015/L13). A.G.-W. was supported by a Life Sciences Research Network Wales (LSRNW) PhD studentship. J.A.D. is supported by a Cancer Research UK Biotherapeutics Drug Discovery Project Award to A.L.P. (project reference C52915/A23946). H. U.-K. was supported by a Cancer Research Wales PhD studentship to A.L.P. L.C. is funded by The HC Roscoe Grant 2016 from the British Medical Association Foundation for Medical Research and by the National Institutes for Health (NIH)/National Institute of Allergy and Infectious Diseases (NIAID) under CEIRS contract HHSN272201400008C. A.L.P. and P.J.R. are funded by Higher Education Funding Council for Wales. The authors acknowledge the Diamond Light Source for beamtime (proposal mx14843) and the staff of beamline I04 for assistance with diffraction data collection. The authors acknowledge Johanne Pentier for technical assistance with SPR data collection, as well as Aaron Wall and Anna Fuller for assistance with FPLC. The authors also acknowledge Elmar Kreiger and the team who maintain the YASARA energy minimisation server, and Pedro Reche and the Immunomedicine Group for maintenance of the Protein Variability Server.

Author contributions

A.T.B. and A.L.P. conceived and designed the study. Modelling of protein–protein interfaces, phylogenetics, protein variability, and interface energy calculations were performed by A.T.B. A.T.B. and A.G.-W. performed crystallisation experiments. A.T.B., A.G.-W., and P.J.R. solved and refined crystallographic structures, and analysed the resultant models. Competition inhibition studies were performed by A.T.B. with advice from L.C. A.T.B. and D.K.C. performed SPR experiments. H.U.-K., J.A.D., and L.C. provided DNA constructs and preliminary data. The manuscript was prepared by A.T.B. and A.L.P.; all other authors reviewed, edited, and approved the manuscript. The study was supervised by A.L.P.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-019-08599-y>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Journal peer review information: *Nature Communications* thanks Pascal Fender and the other anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© Crown 2019

VIROLOGY

Human adenovirus type 26 uses sialic acid–bearing glycans as a primary cell entry receptor

Alexander T. Baker¹, Rosie M. Mundy¹, James A. Davies¹, Pierre J. Rizkallah², Alan L. Parker^{1*}

Adenoviruses are clinically important agents. They cause respiratory distress, gastroenteritis, and epidemic keratoconjunctivitis. As non-enveloped, double-stranded DNA viruses, they are easily manipulated, making them popular vectors for therapeutic applications, including vaccines. Species D adenovirus type 26 (HAdV-D26) is both a cause of EKC and other diseases and a promising vaccine vector. HAdV-D26–derived vaccines are under investigation as protective platforms against HIV, Zika, and respiratory syncytial virus infections and are in phase 3 clinical trials for Ebola. We recently demonstrated that HAdV-D26 does not use CD46 or Desmoglein-2 as entry receptors, while the putative interaction with coxsackie and adenovirus receptor is low affinity and unlikely to represent the primary cell receptor. Here, we establish sialic acid as a primary entry receptor used by HAdV-D26. We demonstrate that 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 important, both as human pathogens and as platforms for therapeutic applications. As pathogens, human adenoviruses (HAdVs) have been isolated from severe infections in both immunocompromised and healthy individuals (1). Some adenoviruses have been associated with acute infection of the eye (2), respiratory (3), and gastrointestinal tract (1). In rare cases, infections prove fatal, as observed in a recent neonatal infection of species D adenovirus type 56 (HAdV-D56) (4), among adult patients in New Jersey with HAdV-B7d (5) and infamously with HAdV-E4, where large epidemics of adenovirus infection have been seen in military recruits (6). However, most fatal infections are observed in immunocompromised individuals (7), such as those suffering from graft-versus-host disease (8) or HIV (9, 10).

Adenoviruses are classified into seven species (A to G) and between 57 and 90 types depending on the taxonomic definitions used (11, 12). Some adenoviruses have been studied in detail, having well-defined receptor tropisms, including as coxsackie and adenovirus receptor (CAR) (13), cluster of differentiation 46 (CD46, also known as membrane cofactor protein) (14), desmoglein 2 (DSG2) (15), or sialic acid–bearing glycans (16). However, most types have low seroprevalence in the population (17), though this varies significantly by geographical location (18, 19). Their rarity means that many types 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 57 canonical types (12).

Species D adenoviruses are associated with several pathogenicities. HAdV-D56 is a potentially fatal emergent respiratory pathogen composed of a recombination between four species D adenoviruses (4). Opportunistic adenovirus infection isolated from patients with HIV/AIDS are most commonly from species D, where they are associated with prolonged shedding in the gastrointestinal tract (9). HAdV-D has also been associated with genital disease (20). The species D adenoviruses are best known, however, for causing epidemic

keratoconjunctivitis (EKC) infections, which is endemic, but not isolated, to Japan (21). Classically, the primary EKC causing adenoviruses have been HAdV-D8, HAdV-D37, and HAdV-D64 (22) [previously classified as 19a (23)]. More recently, other species D adenoviruses have been associated with EKC, including HAdV-D53 (previously classified as HAdV-D22/H8) (24, 25), HAdV-D54 (26), and HAdV-D56 (27).

The double-stranded adenovirus DNA genome makes them readily amenable to genetic modification (28) and therefore has made them attractive candidates for genetic manipulation for therapeutic applications in cancer (oncolytic viruses) (29) and as vaccine vectors (30). 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 (17), makes them attractive platforms for vaccines, as evidenced by their progression through clinical trials for HIV (31), Zika (32), and Ebola treatment (33). However, there remains a lack of understanding regarding their basic biology and mechanisms of cellular infection. This is exemplified by HAdV-D26, which is being investigated as a vaccine vector for Zika (32), HIV (34), and respiratory syncytial virus (35), and has entered phase 3 clinical trials as an Ebola vaccine (33).

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 (36). 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 (36). The deduced low affinity of the interaction between CAR and HAdV-D26 fiber knob makes it unlikely that CAR represents the definitive primary receptor of HAdV-D26.

Here, we demonstrate that HAdV-D26 uses sialic acid–bearing glycans as a primary entry receptor and that this interaction can form a productive infection. We solve the structure of HAdV-D26K in complex with sialic acid [*N*-acetylneuraminic acid (Neu5Ac)], revealing a similar topology to the known sialic acid–interacting

Copyright © 2019
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
License 4.0 (CC BY).

¹Division of Cancer and Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK. ²Division of Infection and Immunity, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK.

*Corresponding author. Email: parkeral@cardiff.ac.uk

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 types.

RESULTS

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

Our recent findings highlight that HAdV-D26K is unlikely to use CD46 as a primary cellular receptor. We further confirmed this observation using surface plasmon resonance to confirm that the HAdV-D26 fiber knob protein does not stably interact with either CD46 or DSG2. A weak affinity for CAR was observed for HAdV-D26K; however, this is unlikely to represent a primary receptor. Previous amino acid sequence alignments demonstrated little conservation of sialic acid-binding residues with the fiber knob domains of the known sialic acid using HAdV-G52SFK (short fiber knob) or canine adenovirus type-2 (CAV-2) (36). However, these alignments indicated that 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 that of HAdV-D8 (Fig. 1A). These three viruses have been shown to cause 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 whether a similar binding mechanism was possible.

These sialic acid-binding viruses all have highly negative predicted isoelectric points (pIs) (Fig. 1A). We calculated the surface electrostatic potentials of these fiber knob proteins at pH 7.35 to simulate the pH of extracellular fluid, using previously published crystal structures where available. There is no published structure of HAdV-D8K; thus, we generated a homology model based on the closest known relative with a crystal structure (Fig. 1B).

The analyzed fiber knob proteins are highly charged. We observed a concentration of positive charge in the central depression around the threefold axis, which corresponds to the previously reported sialic acid-binding sites (Fig. 1, B to D). 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) (37).

HAdV-D26K has a lower predicted pI, 6.49, and a less positive surface potential (Fig. 1, A and 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. 1E).

HAdV-D26 requires cell surface sialic acid for efficient infection

Sequence alignment of HAdV-D26K with these known sialic acid-using viruses, bearing a positively charged apex, showed conservation

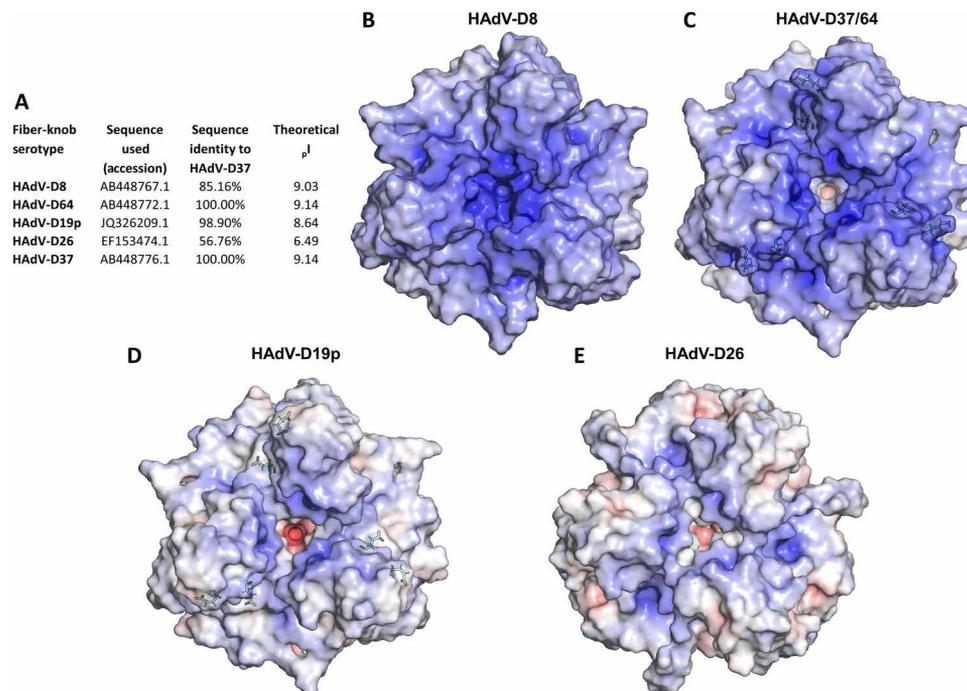


Fig. 1. HAdV-D26K forms a local basic area in the apical depression to facilitate sialic acid binding despite an overall acidic predicted isoelectric point. (A) HAdV-D26K has low (56.76%) sequence identity with fiber knobs known to bind sialic acid by a similar mechanism and an acidic isoelectric point. 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 threefold axis. (E) HAdV-D26 fiber knob is less basic overall but maintains positive potential in the central depression. Surfaces are displayed at ± 10 mV, and the two residues that differ between HAdV-D19p and HAdV-D37/64 are shown as green sticks.

of key binding residues between types (Fig. 2A). We observe complete conservation of Tyr³¹⁴ and Lys³⁴⁹ across the four types and conservation of Asp³¹² with HAdV-D8 (Fig. 2A). Furthermore, while HAdV-D26K Tyr³²⁰ is not conserved, inspection of the crystal structure of HAdV-D37K and HAdV-D19p in complex with sialic acid [Protein Data Bank (PDB) 1UXA and 1UXB, respectively] (37) 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 use sialic acid as a cell entry receptor, we used a replication-incompetent HAdV-C5 vector pseudotyped with the HAdV-D26 fiber knob, expressing a green fluorescent protein (GFP) transgene. We performed infectivity studies in three cell lines, with and without pretreatment with neuraminidase, to remove cell surface sialic acid. The tested cell lines could be infected by the CD46-mediated (Fig. 2B) or CAR-mediated (Fig. 2C) 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 treatment in some cases, an effect that has been previously observed (38, 39). This has been suggested to be due to a reduction in the electrostatic repulsion of the negatively charged capsid of HAdV-C5.

Infection by the HAdV-C5/D26K pseudotype was significantly reduced in all three cell lines following treatment with neuraminidase

(Fig. 2D). This inhibition is significant ($P < 0.005$), resulting in >5-fold decrease in infection in all three cell lines tested, similar to the decrease observed when performing similar experiments with EKC-causing viruses HAdV-D37/53/64, of five- to eightfold inhibition (40). These data indicate that HAdV-C5/D26K is using sialic acid, not CD46 or CAR, to infect these cells.

HAdV-D26 forms a stable complex with sialic acid

We crystallized HAdV-D26K in complex with sialic acid to clarify the mechanism of interaction. Refinement of structures generated from HAdV-D26K crystals soaked in sialic acid shows electron density for a small-molecule ligand in the apical depression (Fig. 3A); this is best described by a racemic mixture of α and β anomers, in conjunction with double conformations of sialic acid (Fig. 3B). The cubic space group (table S1) enabled assembly of the biological trimer. We observed three copies of sialic acid bound within the apex of the fiber knob trimer (Fig. 3C), as previously observed in HAdV-D37 and HAdV-D19p.

Sialic acid binding was observed in structures crystallized at both pH 8.0 (PDB 6QU6) and pH 4.0 (PDB 6QU8). Observation of sialic acid density at high σ values suggests a highly stable interaction (fig. S1). Electron density demonstrates the C2 carboxyl and OH groups in two conformations, and the C6 glycerol group is flexible, with the C7-C8 bond rotating to alter the orientation of the glycerol arm,

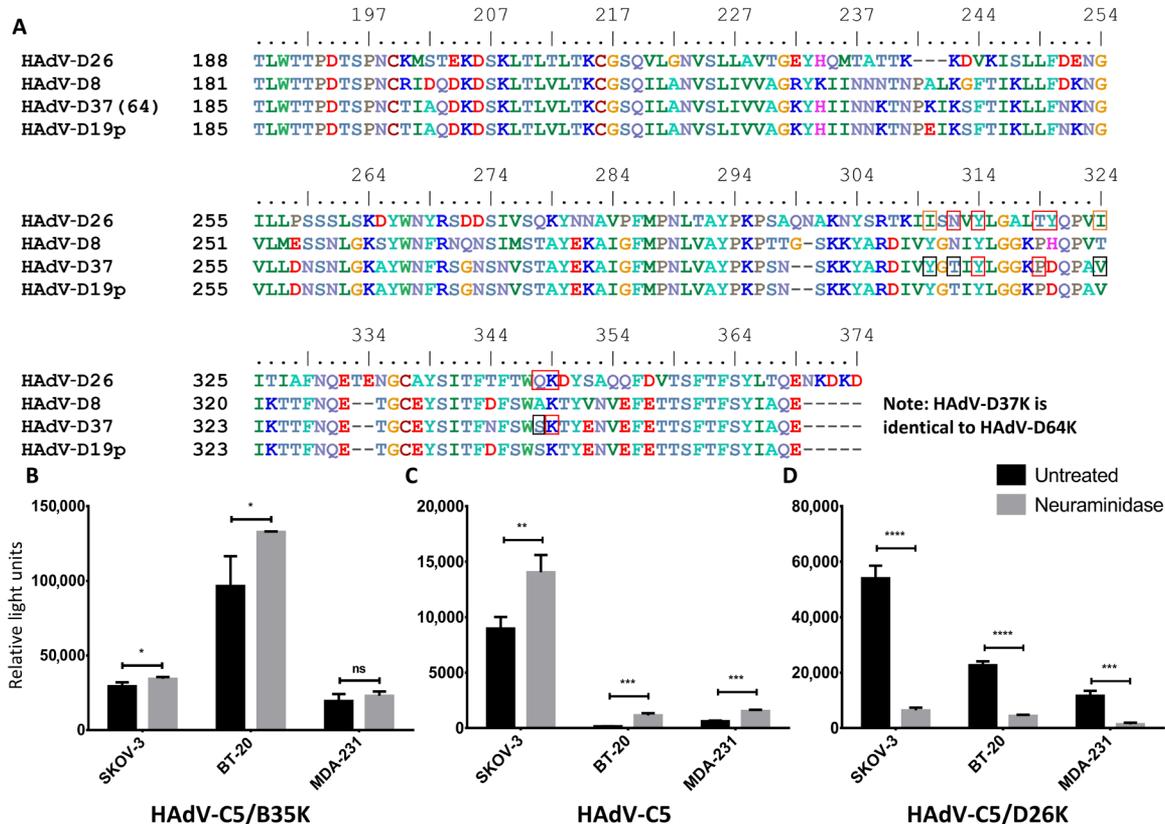


Fig. 2. HAdV-D26K shares key binding residues with sialic acid-using adenoviruses and exploits sialic acid to infect cells. (A) Sequence alignment of HAdV-D26K shows conservation of key binding residues with known sialic acid-binding adenoviruses; top numbering is according to HAdV-D26K. Residues boxed in red form polar contacts with sialic acid, those boxed in black denote contact sialic acid via water bridge, and those boxed in orange indicate hydrophobic contacts; all HAdV-D26K polar contacts also form water bridges. 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 bars indicate \pm SD.

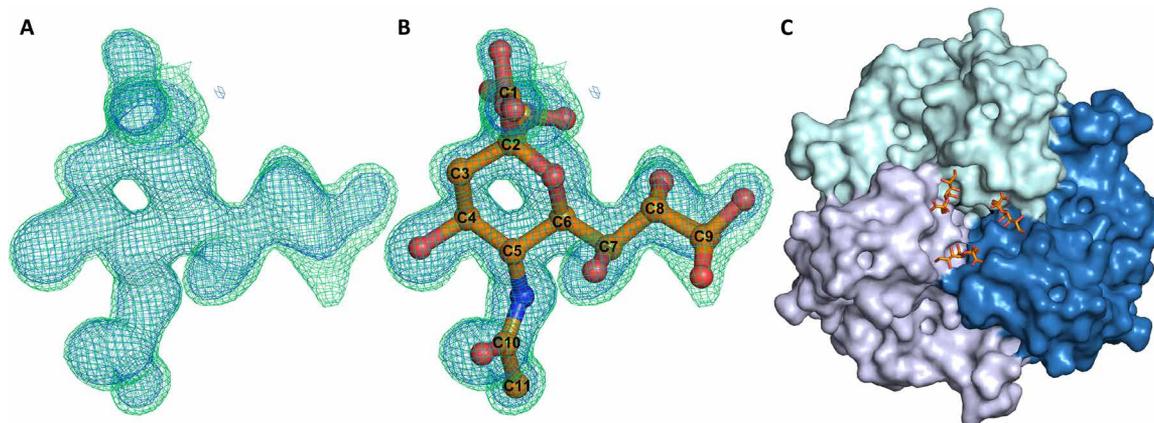


Fig. 3. Sialic acid binds in the apical depression of HAdV-D26 fiber knob protein. The map shows clear density for a ligand (A), which is best described by a double conformer of sialic acid (B). (C) Sialic acid (orange) is seen to bind in three locations in the apical depression of the HAdV-D26 fiber knob, bridging between monomers (shades of blue) of the trimeric assembly. Crystallization statistics are provided in table S1; $2F_o - F_c$ map (blue mesh, $\sigma = 1.5$) and $F_o - F_c$ (green mesh, $\sigma = 3.0$).

relative to the pyranose ring and binding pocket (Fig. 3, A and B, and fig. S1). The glycerol group exhibits further flexibility at the C8–C9 bond, making the terminal oxygen mobile. The distribution of the density for the glycerol group is different at each pH (fig. S1), suggesting that pH could affect the preferred mode of interaction.

The most biologically relevant sialic acid conformation places the carboxyl group axial to the chair conformation pyranose ring (fig. S2), leaving the OH group pointing away from the fiber knob and free to form an $\alpha(2)$ -glycosidic bond as part of a glycan. This is suggestive of a terminal sialic acid residue, as the chain can extend out of the central depression, as was observed in the previously described HAdV-D37K:GD1a glycan structure (37, 41).

HAdV-D26 has a sophisticated sialic acid-binding pocket

Comparison between the HAdV-D26K and HAdV-D37K, the best described of the sialic acid-binding adenoviruses, reveals that several sialic acid contacts are conserved (Fig. 4, A and B). Lys³⁴⁹ and Tyr³¹⁴ are identical, and while Lys³⁴⁹ exhibits some flexibility, all observed lysine conformations form a contact with the carboxyl group of the sialic acid (fig. S3). While 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; hence, 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. 4, A and 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, a pair of water bridges is used.

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 three 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. Similar to 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 (fig. S2), which can be interpreted as

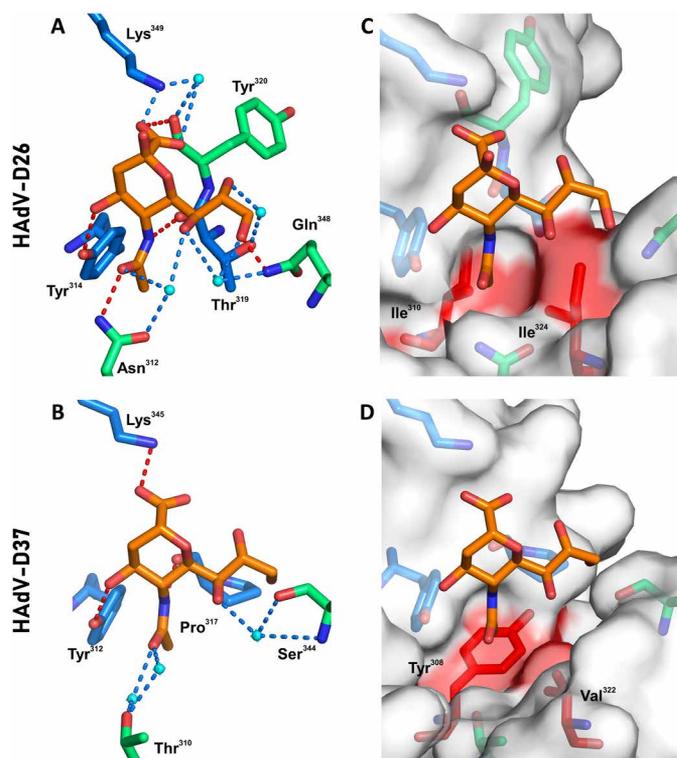


Fig. 4. HAdV-D26K forms a complex interaction network of hydrophobic and electrostatic interactions with sialic acid. Sialic acid (orange) is seen to bind HAdV-D26 (A) and HAdV-D37 (B) through a network of polar contacts (red dashes) and hydrogen bonds (blue dashes). The interaction is stabilized by hydrophobic interactions (red regions on white surface) with the *N*-acetyl CH₃ 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, and other residues are shown as green sticks. Oxygen and nitrogen are seen in red and blue, respectively.

flexibility. However, we suggest that, in HAdV-D26, this is unlikely since it is so well coordinated in all conformations observed, at both pH 8.0 and pH 4.0 (fig. S3). 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-D26 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 pH 8.0 (Fig. 5A) and pH 4.0 (Fig. 5B). While 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 pH 8.0 structure (approximately 0.7), while at pH 4.0, 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 been shown to interact with sialic acid. These include CAV-2 (42), Turkey adenovirus 3 (43), and HAdV-G52 short fiber knob (44, 45), but these viruses interact with sialic acid in lateral regions of the fiber knob, dissimilar from HAdV-D26K. Four other HAdV fiber knob proteins (HAdV-D8/19p/37/64K) have been previously shown to use 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 (Figs. 1A and 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 utilization (46). 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,

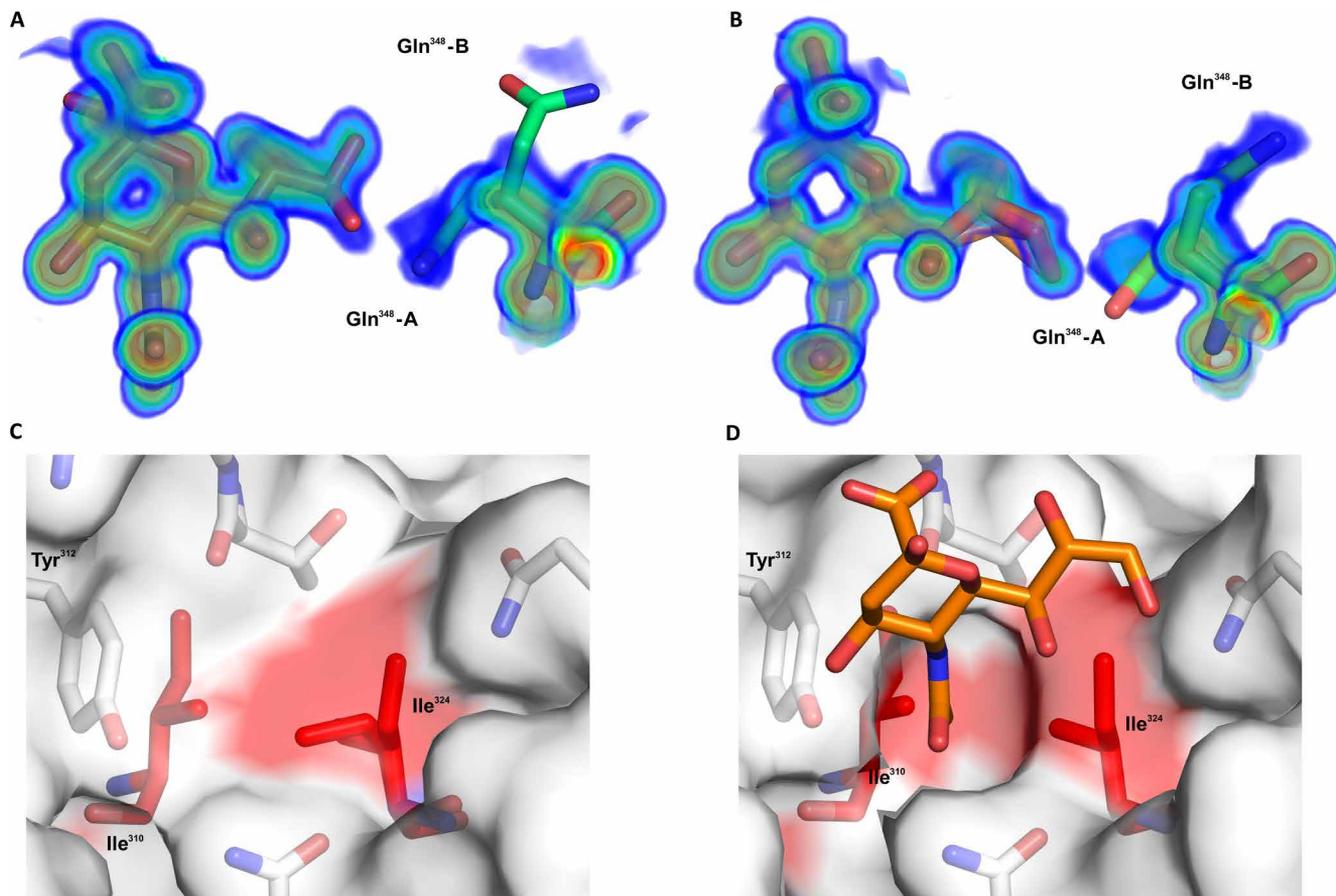


Fig. 5. HAdV-D26K uses an induced-fit mechanism in sialic acid binding. HAdV-D26K residue Gln³⁴⁸ can occupy multiple conformations, with a greater preference for conformation A (capable of forming a polar contact with the glycerol arm of sialic acid) at pH 8.0 (A) than at pH 4.0 (B). (C) Ile³²⁴ has two conformations when HAdV-D26K is unliganded (PDB 6FJO). (D) However, upon sialic acid binding, the Ile³²⁴ adopts a single conformation, creating a hydrophobic indentation around the *N*-acetyl methyl group bounded by Ile³²⁴, Ile³¹⁰, and the ring of Tyr³¹².

we observe a similar electrostatic profile in the modeled fiber knob as seen in HAdV-D37/64 (Fig. 1, B and C). HAdV-D64 has an identical fiber knob domain to that of HAdV-D37; thus, fiber knob interactions with sialic acid are likely to be conserved between these types. 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 acid-binding pocket of HAdV-D26K reveals a much more complex mechanism of interaction than that previously reported for HAdV-D37K (Fig. 4) (37). 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 that 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 stabilize the glycerol side chain (Fig. 4A).

This is one of several examples of HAdV-D26K forming additional contacts with sialic acid that are not observed with HAdV-D37/19p. The substitution of the Thr³¹⁰ and Ser³⁴⁴ found in HAdV-D37K for longer charged residues (Asn³¹² and Gln³⁴⁸, respectively) in HAdV-D26K enables direct polar contacts, as well as additional water-bridge contacts. Substitution of Tyr³⁰⁸ and Val³²² for more hydrophobic isoleucine residues in HAdV-D26 (Ile³¹⁰ and Ile³²⁴, respectively) creates a hydrophobic indentation better tailored to fit around the *N*-acetyl methyl group.

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

The double occupancy of Gln³⁴⁸ may indicate a second induced-fit mechanism. We observe two possible conformations of Gln³⁴⁸ (Fig. 5C). While conformation A does not form any contacts, conformation B forms a polar bond and water bridge, with the sialic acid glycerol group. In HAdV-D37K, the glycerol group forms only a water bridge

from C7-OH to Ser³⁴⁴, the spatial equivalent of Gln³⁴⁸. While we have not determined the preferred conformation of the sialic acid glycerol group versus Gln³⁴⁸ conformation, we observed that it can form polar contacts with it regardless, while Gln³⁴⁸ is in conformation B (Fig. 5C).

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

Gln³⁴⁸ has greater occupancy in a sialic acid-binding conformation (conformation A) at pH 8.0, which corresponds more closely to the physiological conditions in which it would encounter at the cell surface (Fig. 5A). At pH 4.0, 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 summarized by three synchronous mechanisms: an *N*-acetyl anchor composed of a polar contact to Asn³¹² stabilized 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; and, last, a network of polar contacts to the carboxyl, C4 oxygen, and nitrogen atoms, which stabilize the pyranose ring.

This interaction in HAdV-D26K is a much more sophisticated binding mechanism compared to HAdV-D19p and HAdV-D37/64. However, the overall pocket topology and several key residues bear similarities. It may be unexpected 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 precedent for this within adenovirus, with recombination events being reported in numerous settings (23, 24).

It has previously been suggested that many of the species D adenoviruses may have dual sialic acid-binding affinity and CAR affinity (37). This has been observed in HAdV-D37/64, CAV-2 (13, 42), and, now, HAdV-D26 (36). 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 (44), which is shown to bind polysialic acid (47). Previous work has proposed that CAR may be a receptor for many, if not all, of the species D adenoviruses with variable affinity (36, 37) and suggests that sialic acid could also be widely used (37). These findings support that assertion, adding another species D adenovirus, with low sequence similarity, to the pool of adenoviruses observed to bind both CAR and sialic acid.

HAdV-D43, HAdV-D27, and HAdV-D28 fiber knobs share high sequence homology with HAdV-D26K, sharing most of the critical binding residues, and/or having structural homologs at those positions (fig. S4). 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 have an asparagine that 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 it may alter affinity (fig. S4).

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 (fig. S4).

The HAdV-D26K surface electrostatics most closely resemble those of HAdV-D19pK. HAdV-D19pK is capable of binding sialic acid (37), and a limited effect is seen on infection of A549 cell binding after neuraminidase treatment to remove cell surface sialic acid (46). HAdV-D19p binding to Chang C (human conjunctival) cells was completely unaffected by neuraminidase treatment, though binding was very low regardless of neuraminidase treatment (38). This inability to bind Chang C cells was shown to depend on a single lysine residue (Lys²⁴⁰) in the apex of the fiber knob, but distant from the sialic acid-binding pocket, creating a more acidic apical region in the lysine's absence (48). 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 (41). However, it seems unlikely that a protein capable of trivalent sialic acid binding is completely specific for GD1a, a disialylated glycan, given the wide range of available glycan motifs that are di- and trisialylated. 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 toward cells with different glycosylation patterns.

The infectivity assays demonstrate sialic acid utilization by HAdV-C5/D26K in multiple cell types (Fig. 2D), suggesting either that HAdV-D26K binds to a glycosylation pattern conserved in all three cell lines or that HAdV-D26K can bind to variable glycosylation motifs. In this assay, a non-sialic acid-using HAdV-C5 virus (Fig. 2C) pseudotyped with the fiber knob domain of HAdV-D26 was used. This pseudotype (and HAdV-C5/B35K) retains the fiber shaft domain of HAdV-C5, which is 212 amino acids longer than the native HAdV-D26 fiber shaft. This could potentially increase the avidity of the virus for its receptor, providing greater range and flexibility than the equivalent short species D fiber shaft. However, it is unlikely to affect the receptor binding that is largely governed by the interaction with the knob portion of the fiber protein.

Previous work by Nestić *et al.* (49) suggested that $\alpha\text{v}\beta 3$ integrin is used by HAdV-D26 during infection and provided evidence for a role for CAR as a low-affinity receptor for HAdV-D26, supporting our previous structural and biological observations (36). Enhanced $\alpha\text{v}\beta 3$ -dependent infection by HAdV-D26 is likely explained by secondary interactions with the penton base, which has a shorter Arginine, Glycine, Aspartic Acid (RGD)-containing loop than HAdV-C5, demonstrated previously using cryo-electron microscopy (50). The infectivity of the pseudotyped HAdV-C5/D26K vector described here is not limited by this reduced integrin affinity due to the retention of the HAdV-C5 penton base protein; thus, conclusions cannot be drawn from the present study regarding secondary interactions that govern internalization of HAdV-D26.

Our 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 (fig. S4). 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 (51) may make effective anti-HAdV-D26 therapies.

MATERIALS AND METHODS

Generation of fiber knob pseudotyped HAdV-C5 viral vectors

HAdV-C5 viruses pseudotyped with the HAdV-D26 or HAdV-D35 fiber knob proteins were generated by the recombineering method, as published by Stanton *et al.* (28). In brief, a marker cassette was generated using the SacB cassette template (28), with the SacB primer pair (table S1) with homology to the HAdV-C5 fiber knob DNA sequence before the Threonine, Leucine, Tryptophan (TLW) hinge sequence (forward primer) and after the stop codon (reverse primer). This template was integrated into a bacterial artificial chromosome (BAC) containing the genome of a GFP-expressing HAdV-C5, which had been rendered replication incompetent by deletion of the E1A gene, now containing a marker cassette instead of the HAdV-C5 fiber knob domain.

The DNA sequence of the HAdV-D26K and HAdV-B35K fiber knobs was amplified using the 26K and 35K primer pairs, respectively (table S1), containing similar homology to the SacB cassette. A second round of recombineering was used to generate the final HAdV-C5 pseudotyped genomes by integrating the HAdV-D26K and HAdV-B35K polymerase chain reaction (PCR) transcripts. After recombineering, the new BACs were sequenced to confirm the correct fiber knob DNA sequence.

The BAC DNA for the new vector genomes was transfected into 293 cell line stably expressing the tetracycline repressor protein (TREx-293) cells (1.5×10^6), cultured in a T25 cell bind flask (Corning) in 5 ml of Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% v/v fetal bovine serum, using the effectene system (QIAGEN). Cells were kept in culture until a cytopathic effect (CPE) was observed, at which point they were harvested by scraping and centrifugation at 1200g for 3 min. Cells were resuspended in 1 ml of media and frozen at -80°C to create a crude stock of virus. TREx-293 cells were cultured at 70% confluency in 5× T150 cell bind flasks (Corning) containing 20 ml of complete DMEM and then inoculated with 10 μl of crude virus stock. Cells were maintained in culture until CPE was observed and harvested by scraping and centrifugation at 1200g.

Virus was then purified from this cell pellet using the Cesium Chloride (CsCl) gradient method (52). Virus titer was determined in viral particles per milliliter (VP/ml) using the Pierce BCA Protein Assay Kit, assuming 4×10^9 VP/ μg of protein (52). By using this method, we were able to generate HAdV-C5 viruses pseudotyped with the fiber knob domains of HAdV-D26 or HAdV-B35. These viruses retained the HAdV-C5 fiber shaft domain and were replication incompetent.

Infectivity assays

Cells were seeded at a density of 30,000 cells per well in a flat-bottomed 96-well cell culture plate and incubated overnight at 37°C to adhere. Cells were washed twice with 200 μ l of phosphate buffered saline (PBS), and 50 μ l of neuraminidase (Sigma-Aldrich, cat. no. 11080725001) was added to the appropriate wells at a concentration of 50 mU/ml, diluted in serum-free media, and incubated for 1 hour at 37°C. Cells were cooled on ice and washed twice with 200 μ l of PBS. GFP-expressing, replication-incompetent viruses were added to the appropriate wells at a concentration of 2000 or 5000 VP per cell, in 100 μ l of serum-free media, at 4°C, and incubated on ice for 1 hour. Serum-free media alone were added to uninfected control wells. Cells were washed twice with 200 μ l of cold PBS, and complete media were added (DMEM, 10% fetal calf serum) and incubated for 48 hours at 37°C. Cells were then trypsinized and transferred to a 96-well V-bottom plate, washed twice in 200 μ l of PBS and fixed in 2% paraformaldehyde for 20 min before wash, and resuspended in 200 μ l of PBS.

Samples were run in triplicate and analyzed by flow cytometry on Attune NxT (Thermo Fisher Scientific), analyzed using FlowJo v10 (FlowJo, LLC), gating sequentially on singlets, cell population, and GFP-positive cells. Levels of infection were described in terms of total fluorescence, defined as the percentage of GFP-positive cells (% + percentage positive) multiplied by the median fluorescence intensity of the GFP-positive population.

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), and 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 European Bioinformatics Institute Clustal Omega tool (53).

Generation of recombinant fiber knob protein

SG13009 *Escherichia coli* harboring pREP-4 plasmid and pQE-30 expression vector containing the fiber knob DNA sequence were cultured in 20 ml of LB broth with ampicillin (100 μ g/ml) and kanamycin (50 μ g/ml) overnight from glycerol stocks made in previous studies (54). One liter of Terrific Broth (modified, Sigma-Aldrich) containing ampicillin (100 μ g/ml) and kanamycin (50 μ g/ml) was inoculated with the overnight *E. coli* culture and incubated at 37°C until they reached an optical density of 0.6. Isopropyl- β -D-thiogalactopyranoside was then added to a final concentration of 0.5 mM, and the culture was incubated at 37°C for 4 hours. Cells were then harvested by centrifugation at 3000g, resuspended in lysis buffer [50 mM tris (pH 8.0), 300 mM NaCl, 1% (v/v) NP-40, lysozyme (1 mg/ml), and 1 mM β -mercaptoethanol], and incubated at room temperature for 30 min. Lysate was clarified by centrifugation at 30,000g for 30 min and filtered through a 0.22- μ m syringe filter (Millipore, Abingdon, UK). Clarified lysate was then loaded onto a 5-ml HisTrap FF nickel affinity chromatography column (GE Healthcare) at 2.0 ml/min and washed with 5 column volumes into elution buffer A [50 mM tris (pH 8.0), 300 mM NaCl, and 1 mM β -mercaptoethanol]. Protein was eluted by 30-min gradient elution from buffer A to B (buffer A + 400 mM imidazole). Fractions were analyzed by reducing SDS-polyacrylamide gel electrophoresis (SDS-PAGE), and fiber knob-containing fractions were further purified using a Superdex 200 10/300 size exclusion chroma-

tography column (GE) in crystallization buffer [10 mM tris (pH 8.0) and 30 mM NaCl]. Fractions were analyzed by SDS-PAGE and pure fractions were concentrated by centrifugation in VivaSpin (10,000 molecular weight cutoff) (Sartorius, Goettingen, Germany) preceding crystallization.

Crystallization and structure determination

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

Crystals of HAdV-D26K appeared in PACT Premier conditions B01 and B04 [0.1 M MIB (malonic acid, imidazole, and boric acid) and 25% w/v polyethylene glycol 1500, pH 4.0 and pH 8.0, respectively], within 1 to 7 days. Crystals were then soaked in reservoir solution containing Neu5Ac (Sigma-Aldrich, cat. no. A2388) at a final concentration of 10 mM and incubated overnight before 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 0.9795 Å. During data collection, crystals were maintained in a cold air stream at 100°K. Dectris PILATUS 6M detectors recorded the diffraction patterns, which were analyzed and reduced with XDS (55), Xia2, DIALS (56), and autoProc (57). Scaling and merging data were completed with Pointless, Aimless and Truncate from the CCP4 package (58). Structures were solved with Phaser, COOT was used to correct the sequences and adjust the models, and REFMAC5 was used to refine the structures and calculate maps. Graphical representations were prepared with PyMOL. Reflection data and final models were deposited in the PDB database with accession codes 6QU6, 6QU8, and 6FJO. Full crystallographic refinement statistics are given in table S1.

Calculation of electrostatic surface potentials and pIs

HAdV-D37, HAdV-D19p, and HAdV-D26 used PDB IUXA (37), PDB IUXB (37), and PDB 6QU8, respectively, as the input. HAdV-D8 was calculated using a homology model, generated as described below, for input.

The PDB2PQR server (V 2.1.1) (http://nbc-222.ucsd.edu/pdb2pqr_2.1.1/) was used to assign charge and radius parameters using the PARSE force field and assign protonation states using Propka, at pH 7.35. APBS was used to calculate electrostatic surface potentials, and the map output was visualized in PyMOL (59).

Homology modeling of adenovirus type 8

The I-TASSER protein structure and function prediction server (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) (60) was used to generate a homology model of HAdV-D8 based on the published sequence of HAdV-D2 (22), using the published structure of its closest relative (by sequence identity), HAdV-D19p (37). The resultant monomer was then copied three times, using the HAdV-19p trimer as a template and the monomers aligned in PyMOL to generate a model of the complete HAdV-D8K trimer.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/5/9/eaax3567/DC1>

Fig. S1. Sialic acid forms a stable interaction with HAdV-D26K 654 at both pH 4.0 (PDB 6QU6) and pH 8.0 (PDB 6QU8).

Fig. S2. Structure of sialic acid (Neu5Ac) in a biologically relevant conformation.

Fig. S3. HAdV-D26K forms a similar interaction with sialic acid at both pH 4.0 (PDB 6QU6) and pH 8.0 (PDB 6QU8) through a combination of polar, water bridge, and hydrophobic interactions.

Fig. S4. Species D adenoviruses conserve known sialic acid-binding residues.

Table S1. Data collection and refinement statistics for structures generated in this study.

Table S2. Primers used to generate recombinering PCR products in this study.

REFERENCES AND NOTES

1. T. Lion, Adenovirus infections in immunocompetent and immunocompromised patients. *Clin. Microbiol. Rev.* **27**, 441–462 (2014).
2. D. Garcia-Zalznak, 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).
3. R. Cviljak, T. tot, A. R. Falsey, E. Huljev, J. Vranes, S. Ljubin-Sternak, Viral pathogens associated with acute respiratory illness in hospitalized adults and elderly from Zagreb, Croatia, 2016 to 2018. *J. Med. Virol.* **91**, 1202–1209 (2019).
4. C. M. Robinson, G. Singh, C. Henquell, M. P. Walsh, H. Peigue-Lafeuille, D. Seto, M. S. Jones, D. W. Dyer, J. Chodosh, Computational analysis and identification of an emergent human adenovirus pathogen implicated in a respiratory fatality. *Virology* **409**, 141–147 (2011).
5. M. E. Killerby, F. Rozwadowski, X. Lu, M. Caulcrick-Grimes, L. McHugh, A. M. Haldeman, T. Fulton, E. Schneider, S. K. Sakhivel, J. Bhatnagar, D. B. Rabeneck, S. Zaki, S. I. Gerber, J. T. Watson, Respiratory illness associated with emergent human adenovirus genome type 7d, New Jersey, 2016–2017. *Open Forum Infect. Dis.* **6**, ofz017 (2019).
6. 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).
7. S. Khanal, P. Ghimire, A. S. Dharmoon, The repertoire of adenovirus in human disease: The innocuous to the deadly. *Biomedicine* **6**, E30 (2018).
8. 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).
9. 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).
10. J. C. Hierholzer, Adenoviruses in the immunocompromised host. *Clin. Microbiol. Rev.* **5**, 262–274 (1992).
11. 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.
12. HAdV Working Group, <http://hadvwwg.gmu.edu/>.
13. 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).
14. A. Gaggar, D. M. Shayakhmetov, A. Lieber, CD46 is a cellular receptor for group B adenoviruses. *Nat. Med.* **9**, 1408–1412 (2003).
15. E. Vassal-Stermann, G. Effantin, C. Zubieta, W. Burmeister, F. Iseni, H. Wang, A. Lieber, G. Schoehn, P. Fender, CryoEM structure of adenovirus type 3 fibre with desmoglein 2 shows an unusual mode of receptor engagement. *Nat. Commun.* **10**, 1181 (2019).
16. 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).
17. P. Abbink, A. A. C. Lemckert, B. A. Ewald, D. M. Lynch, M. Denholtz, S. Smits, L. Holterman, I. Damen, R. Vogels, A. R. Thorne, K. L. O'Brien, A. Carville, K. G. Mansfield, J. Goudsmit, M. J. E. Havenga, D. H. Barouch, Comparative seroprevalence and immunogenicity of six rare serotype recombinant adenovirus vaccine vectors from subgroups B and D. *J. Virol.* **81**, 4654–4663 (2007).
18. M. Pauly, E. Hoppe, L. Mugisha, K. Petrzalkova, C. Akoua-Koffi, E. Couacy-Hymann, A. Anoh, A. Mossoun, G. Schubert, L. Wiersma, S. Pascale, J. J. Muyembe, S. Karhemere, S. Weiss, S. Leendertz, S. Calvignac-Spencer, F. H. Leendertz, B. Ehlers, High prevalence and diversity of species D adenoviruses (HAdV-D) in human populations of four Sub-Saharan countries. *Viol. J.* **11**, 25 (2014).
19. R. S. Dakin, A. L. Parker, C. Delles, S. A. Nicklin, A. H. Baker, Efficient transduction of primary vascular cells by the rare adenovirus serotype 49 vector. *Hum. Gene Ther.* **26**, 312–319 (2015).
20. P. D. Swenson, M. S. Lowens, C. L. Celum, J. C. Hierholzer, Adenovirus types 2, 8, and 37 associated with genital infections in patients attending a sexually transmitted disease clinic. *J. Clin. Microbiol.* **33**, 2728–2731 (1995).
21. K. Aoki, Y. Tagawa, A twenty-one year surveillance of adenoviral conjunctivitis in Sapporo, Japan. *Int. Ophthalmol. Clin.* **42**, 49–54 (2002).
22. H. Kaneko, T. Iida, H. Ishiko, T. Ohguchi, T. Ariga, Y. Tagawa, K. Aoki, S. Ohno, T. Suzutani, Analysis of the complete genome sequence of epidemic keratoconjunctivitis-related human adenovirus type 8, 19, 37 and a novel serotype. *J. Gen. Virol.* **90**, 1471–1476 (2009).
23. X. Zhou, C. M. Robinson, J. Rajaiya, S. Dehghan, D. Seto, M. S. Jones, D. W. Dyer, J. Chodosh, Analysis of human adenovirus type 19 associated with epidemic keratoconjunctivitis and its reclassification as adenovirus type 64. *Invest. Ophthalmol. Vis. Sci.* **53**, 2804–2811 (2012).
24. M. P. Walsh, A. Chintakuntlawar, C. M. Robinson, I. Madisch, B. Harrach, N. R. Hudson, D. Schnurr, A. Heim, J. Chodosh, D. Seto, M. S. Jones, Evidence of molecular evolution driven by recombination events influencing tropism in a novel human adenovirus that causes epidemic keratoconjunctivitis. *PLOS ONE* **4**, e5635 (2009).
25. K. Aoki, N. Kitaichi, R. Hinokuma, S. Ohno, Characteristics of epidemic keratoconjunctivitis due to novel adenovirus types. *J. Clin. Ophthalmol. Eye Disord.* **1**, 1013 (2017).
26. K. Matsuura, Y. Terasaka, E. Uchio, Y. Saeki, T. Fujimoto, N. Hanaoka, D. Miyazaki, Y. Inoue, Human adenoviral type 54 keratoconjunctivitis accompanied by stellate keratitis and keratic precipitates: Two cases. *BMC Ophthalmol.* **19**, 7 (2019).
27. G. Huang, W. Yao, W. Yu, L. Mao, H. Sun, W. Yao, J. Tian, L. Wang, Z. Bo, Z. Zhu, Y. Zhang, Z. Zhao, W. Xu, Outbreak of epidemic keratoconjunctivitis caused by human adenovirus type 56, China, 2012. *PLOS ONE* **9**, e110781 (2014).
28. 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, 664–662, 668 (2008).
29. H. Uusi-Kerttula, S. Hulin-Curtis, J. Davies, A. L. Parker, Oncolytic adenovirus: Strategies and insights for vector design and immuno-oncology applications. *Viruses* **7**, 6009–6042 (2015).
30. D. Majhen, N. Calderon, N. Chandra, C. A. Fajardo, A. Rajan, R. Alemany, J. Custers, Adenovirus-based vaccines for fighting infectious diseases and cancer: Progress in the field. *Hum. Gene Ther.* **25**, 301–317 (2014).
31. D. H. Barouch, F. L. Tomaka, F. Wegmann, D. J. Stieh, G. Alter, M. L. Robb, N. L. Michael, L. Peter, J. P. Nkolola, E. N. Borducchi, A. Chandrashekar, D. Jetton, K. E. Stephenson, W. Li, B. Korber, G. D. Tomaras, D. C. Montefiori, G. Gray, N. Frahm, M. J. McElrath, L. Baden, J. Johnson, J. Hutter, E. Swann, E. Karita, H. Kibuuka, J. Mpendo, N. Garrett, K. Mngadi, K. Chinyenze, F. Priddy, E. Lazarus, F. Laher, S. Nitayapan, P. Pitisuttithum, S. Bart, T. Campbell, R. Feldman, G. Luckinger, C. Borremans, K. Callewaert, R. Roten, J. Sadoff, L. Scheppeler, M. Weijts, K. Feddes-de Boer, D. van Manen, J. Vreugdenhil, R. Zahn, L. Lavreys, S. Nijjs, J. Tolboom, J. Hendriks, Z. Euler, M. G. Pau, H. Schuitemaker, 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). *Lancet* **392**, 232–243 (2018).
32. F. Cox, L. van der Fits, P. Abbink, R. A. Larocca, E. van Huizen, E. Saeland, J. Verhagen, R. Peterson, J. Tolboom, B. Kaufmann, H. Schuitemaker, D. H. Barouch, R. Zahn, 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).
33. 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, <https://clinicaltrials.gov/ct2/show/NCT02661464?term=adenovirus&recrs=a&cond=ebola&rank=1>.
34. Trial of the Safety and Immunogenicity of an Oral, Replicating Ad26 Vectorized HIV-1 Vaccine—Full Text View—ClinicalTrials.gov, <https://clinicaltrials.gov/ct2/show/NCT02366013>.
35. 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, <https://clinicaltrials.gov/ct2/show/NCT03795441>.
36. A. T. Baker, A. Greenshields-Watson, L. Coughlan, J. A. Davies, H. Uusi-Kerttula, D. K. Cole, P. J. Rizkallah, A. L. Parker, Diversity within the adenovirus fiber knob hypervariable loops influences primary receptor interactions. *Nat. Commun.* **10**, 741 (2019).
37. 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).
38. 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).
39. S. M. Arcasoy, J. Latoche, M. Gondor, S. C. Watkins, R. A. Henderson, R. Hughey, O. J. Finn, J. M. Pilewski, MUC1 and other sialoglycoconjugates inhibit adenovirus-mediated gene transfer to epithelial cells. *Am. J. Respir. Cell Mol. Biol.* **17**, 422–435 (1997).
40. N. Chandra, L. Frångsmyr, S. Imhof, R. Caraballo, M. Elofsson, N. Arnberg, Sialic acid-containing glycans as cellular receptors for ocular human adenoviruses: Implications for tropism and treatment. *Viruses* **11**, E395 (2019).
41. E. C. Nilsson, R. J. Storm, J. Bauer, S. M. C. Johansson, A. Lookene, J. Ångström, M. Hedenström, T. L. Eriksson, L. Frångsmyr, S. Rinaldi, H. J. Willison, F. P. Domellöf, T. Stehle, N. Arnberg, The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. *Nat. Med.* **17**, 105–109 (2011).
42. E. Seiradake, D. Henaff, H. Wodrich, O. Billet, M. Perreau, C. Hippert, F. Mennechet, G. Schoehn, H. Lortat-Jacob, H. Dreja, S. Ibanes, V. Kalatzis, J. P. Wang, R. W. Finberg,

- S. Cusack, E. J. Kremer, The cell adhesion molecule “CAR” and sialic acid on human erythrocytes influence adenovirus in vivo biodistribution. *PLOS Pathog.* **5**, e1000277 (2009).
43. A. K. Singh, M. A. Berbis, M. Z. Ballmann, M. Kilcoyne, M. Menéndez, T. H. Nguyen, L. Joshi, F. J. Cañada, J. Jiménez-Barbero, M. Benkő, B. Harrach, M. J. van Raaij, Structure and sialyllactose binding of the carboxy-terminal head domain of the fibre from a siadenovirus, turkey adenovirus 3. *PLOS ONE* **10**, e0139339 (2015).
 44. A. Lenman, A. M. Liaci, Y. Liu, C. Årdahl, A. Rajan, E. Nilsson, W. Bradford, L. Kaeshammer, M. S. Jones, L. Frängsmyr, T. Feizi, T. Stehle, N. Arnberg, Human adenovirus 52 uses sialic acid-containing glycoproteins and the coxsackie and adenovirus receptor for binding to target cells. *PLOS Pathog.* **11**, e1004657 (2015).
 45. M.-C. Amoureux, B. Coulibaly, O. Chinot, A. Loundou, P. Metellus, G. Rougon, D. Figarella-Branger, 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).
 46. N. Arnberg, A. H. Kidd, K. Edlund, F. Olfat, G. Wadell, Initial interactions of subgenus D adenoviruses with A549 cellular receptors: Sialic acid versus α_v integrins. *J. Virol.* **74**, 7691–7693 (2000).
 47. A. Lenman, A. M. Liaci, Y. Liu, L. Frängsmyr, M. Frank, B. S. Blaum, W. Chai, I. I. Podgorski, B. Harrach, M. Benkő, T. Feizi, T. Stehle, N. Arnberg, Polysialic acid is a cellular receptor for human adenovirus 52. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4264–E4273 (2018).
 48. 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).
 49. D. Nestić, T. G. Uil, J. Ma, S. Roy, J. Vellinga, A. H. Baker, J. Custers, D. Majhen, $\alpha\beta 3$ integrin is required for efficient infection of epithelial cells with human adenovirus type 26. *J. Virol.* **93**, e01474-18 (2019).
 50. X. Yu, D. Veesler, M. G. Campbell, M. E. Barry, F. J. Asturias, M. A. Barry, V. S. Reddy, Cryo-EM structure of human adenovirus D26 reveals the conservation of structural organization among human adenoviruses. *Sci. Adv.* **3**, e1602670 (2017).
 51. S. Spjut, W. Qian, J. Bauer, R. Storm, L. Frängsmyr, T. Stehle, N. Arnberg, M. Elofsson, A potent trivalent sialic acid inhibitor of adenovirus type 37 infection of human corneal cells. *Angew. Chem. Int. Ed. Engl.* **50**, 6519–6521 (2011).
 52. H. Uusi-Kerttula, M. Legut, J. Davies, R. Jones, E. Hudson, L. Hanna, R. J. Stanton, J. D. Chester, A. L. Parker, Incorporation of peptides targeting EGFR and FGFR1 into the adenoviral fiber knob domain and their evaluation as targeted cancer therapies. *Hum. Gene Ther.* **26**, 320–329 (2015).
 53. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, D. G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
 54. H. Uusi-Kerttula, J. Davies, L. Coughlan, S. Hulin-Curtis, R. Jones, L. Hanna, J. D. Chester, A. L. Parker, Pseudotyped $\alpha\beta 6$ integrin-targeted adenovirus vectors for ovarian cancer therapies. *Oncotarget* **7**, 27926–27937 (2016).
 55. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
 56. G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter, G. Evans, DIALS: Implementation and evaluation of a new integration package. *Acta Crystallogr. D Struct. Biol.* **74**, 85–97 (2018).
 57. C. Vornrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 293–302 (2011).
 58. E. J. Dodson, M. Winn, A. Ralph, Collaborative computational project, number 4: Providing programs for protein crystallography. *Methods Enzymol.* **277**, 620–633 (1997).
 59. 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–W667 (2004).
 60. J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).

Acknowledgments: We acknowledge the Diamond Light Source for beamtime, funded under grant number MX14843, and the staff of beamline I04 for assistance with diffraction data collection. **Funding:** A.T.B. was supported by a Tenovus Cancer Care PhD studentship to A.L.P. (project reference PhD2015/L13) and the Ludo Frevet Crystallography Scholarship from the International Centre for Diffraction Data (ICDD). J.A.D. was supported by a Cancer Research UK Biotherapeutics Drug Discovery Project Award to A.L.P. (project reference C52915/A23946). R.M.M. was supported by the Cardiff Universities Research Opportunities Placement (CUROP). A.L.P. and P.J.R. were funded by the Higher Education Funding Council for Wales. **Author contributions:** A.T.B. and A.L.P. conceived the study, which was designed in collaboration with P.J.R. Sequence alignments, modeling of electrostatic surfaces, and neuraminidase cleavage cell infectivity assays were performed by A.T.B. The strategy to generate the pseudotyped vectors HAdV-C5/D26K and HAdV-C5/B35K by recombineering was designed by J.A.D. who generated the vectors with A.T.B. Crystallization of HAdV-D26 fiber knob protein was performed by A.T.B., and soaking experiments were performed by A.T.B. and R.M.M. Diffraction data were collected by A.T.B., R.M.M., and P.J.R. Structures were solved, refined, and analyzed by R.M.M. and A.T.B. The study was supervised by A.L.P. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Structural data are available from the worldwide Protein Data Bank, and sequence data are available from the National Center for Biotechnology Information (NCBI). Additional data related to this paper may be requested from the authors.

Submitted 16 March 2019
 Accepted 29 July 2019
 Published 4 September 2019
 10.1126/sciadv.aax3567

Citation: A. T. Baker, R. M. Mundy, J. A. Davies, P. J. Rizkallah, A. L. Parker, Human adenovirus type 26 uses sialic acid-bearing glycans as a primary cell entry receptor. *Sci. Adv.* **5**, eaax3567 (2019).

Human adenovirus type 26 uses sialic acid–bearing glycans as a primary cell entry receptor

Alexander T. Baker, Rosie M. Mundy, James A. Davies, Pierre J. Rizkallah and Alan L. Parker

Sci Adv 5 (9), eaax3567.

DOI: 10.1126/sciadv.aax3567

ARTICLE TOOLS

<http://advances.sciencemag.org/content/5/9/eaax3567>

SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2019/08/30/5.9.eaax3567.DC1>

REFERENCES

This article cites 55 articles, 15 of which you can access for free
<http://advances.sciencemag.org/content/5/9/eaax3567#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science Advances* is a registered trademark of AAAS.