Capsid protein functions of enteric human adenoviruses

Anandi Rajan
Capsid protein functions of enteric human adenoviruses

Anandi Rajan

Department of Clinical Microbiology/Unit of Virology
Umeå 2018
To Amma, Papa and Chechi, without whom I don’t know what I would do!
# Table of Contents

## TABLE OF CONTENTS
1

## ABSTRACT
III

## POPULAR SCIENCE SUMMARY
V

## POPULÄRVETENSKAPLIG SAMMANFATTNING
VII

## ABBREVIATIONS
IX

## AIMS OF THE THESIS
XII

## LIST OF PUBLICATIONS
XIII

## HUMAN ADENOVIRUSES
1

### Background
1
  - History
  - Taxonomy and Diversity

### Epidemiology and Clinical Features
3

## Structural Proteins
5
  - Hexon (pII)
  - Penton base (pIII)
  - Fiber (pIV)
  - Minor Capsid Proteins
    - Protein VI (pVI)
    - Protein IIIA (pIIIA)
    - Protein VIII (pVIII)
    - Protein IX (pIX)
  - Core Proteins


Abstract

Human adenoviruses cause respiratory illnesses, epidemic conjunctivitis and infantile gastroenteritis. There are around 85 known types of human adenoviruses. Human adenovirus types 40 and 41 cause enteric infections in infants worldwide and viral gastroenteritis caused by these viruses is third only to infections caused by rotaviruses and caliciviruses.

Human adenoviruses use various receptors for attachment onto the different host cells. Coxsackievirus and adenovirus receptor, CD46, sialic acid, coagulation factors IX and X, lactoferrin and heparan sulfate are some receptors and molecules which the hexon and fiber proteins (components of the capsid) bind for direct or indirect cellular attachment. The penton base protein (another component of the capsid) is responsible for the internalization of the virus into the host cell. An arginine-glycine-aspartic acid amino acid motif is present in most but not all adenovirus penton base proteins and mediates interaction with αv integrins, resulting in internalization.

The enteric human adenoviruses are unique since they do not have this arginine-glycine-aspartic acid motif on their penton base. However, they do contain amino acid motifs that bind to other integrins that are present in the human gastrointestinal tract. Using a library of hamster cells expressing specific human integrins, along with recombinant soluble penton bases and commercially available soluble laminins, we identified laminin-binding integrins as co-receptors for entry and infection of human intestinal HT-29 cells by the enteric human adenoviruses.

Human adenovirus types 40, 41 and 52 are the only three human adenoviruses that have two different fiber proteins, one long and one short. By performing cell binding and infection experiments, we have found that the receptor for the short fiber of HAdV-52 is sialic acid-containing glycans and the long fiber receptor is CAR and that most of the binding was dependent on sialic acid-containing glycans. We also observed that the short fiber of human adenovirus type 40
interacts with soluble heparin or cell surface heparan sulfate. Further investigation pointed out that the specific sulfate groups on heparin/heparan sulfate (sulfated glycosaminoglycans) are important for this binding. Also, we identified that the interaction and utilization of these glycosaminoglycans as receptors is dependent on exposure to low pH.

We also studied the potential mechanism behind the symptoms caused by these enteric adenoviruses in enteroendocrine cells called enterochromaffin cells. We could show that the short fiber and the hexon of human adenovirus type 41 stimulated release of serotonin from the enterochromaffin cells, which can be a cause of vomiting and diarrhea.

These studies have given us insight into the role of enteric human adenovirus capsid proteins as ligands to hitherto unidentified receptors and co-receptors. We also show that these molecules play important functions in the virus’ infectious cycle and probably also in their disease mechanism of host cells.
Popular science summary

Adenoviruses are viruses that can infect diverse organisms ranging from reptiles, fish and birds to mammals. In this thesis, I focus on adenoviruses that infect humans. There are multiple types of human adenoviruses that cause different diseases, such as upper respiratory tract, eye or stomach infections. My studies mainly focus on two human adenoviruses (types 40 and 41) that cause stomach or gastrointestinal infections. They primarily cause disease in infants and are the third most frequent cause of viral gastroenteritis in the world.

Human adenoviruses carry information for their replication within a DNA molecule. This DNA is encased within a protein structure called a capsid. The most abundant capsid protein is the hexon followed by the penton base. Another capsid protein structure called the fiber protrudes from these penton bases. For a human adenovirus to infect a cell and cause disease, it has to first attach to the cellular surface and then enter the cell, where it multiplies. The fiber protein carries out the first step of infection, where the virus attaches to the host cell. The fiber protein interacts with a molecule on the cell surface called a primary receptor. The second step of the infectious cycle is cell entry, which is initiated when the penton base protein interacts with another cell surface molecule called the secondary receptor.

What is interesting is that human adenovirus types 40, 41 and 52 are different from the other human adenoviruses because they possess two different fiber proteins called a long fiber and a short fiber (the other types only have one). Types 40 and 41 also have a different penton base. The role of the short fiber and penton base proteins in the infectious cycle of human adenovirus types 40 and 41 has been an enigma. In this study, we investigated the functions of the short fiber and the penton base of human adenovirus types 40 and 41. We were also successful in identifying cellular receptors for the long and short fibers of the recently isolated human adenovirus 52.
In the first study, we identified novel integrins, such as $\alpha_6\beta_1$ and $\alpha_6\beta_4$ that interact with the penton base of types 40 and 41. Integrins are cell surface protein molecules that consist of a paired $\alpha$ subunit and a $\beta$ subunit. We performed binding and infection experiments on intestinal cells and evaluated the interaction of the penton base to the integrins. We used similar approaches and methods to discover the receptor molecule for the short fiber protein of types 40 and 41, which is a protein molecule containing extended sugar chains called heparan sulfate. For type 52, the long fiber interacted with a protein called CAR (coxsackievirus and adenovirus receptor) and the short fiber interacted with long chains of sialic acid. These viruses use these identified receptors to attach to and enter the cell for infection.

Next, we investigated the effect of enteric adenoviruses and their capsid proteins on the primary serotonin secreting cells called the enterochromaffin cells located in the small intestine. Serotonin, a signaling molecule, is involved in several gastrointestinal disorders and symptoms such as nausea, vomiting, intestinal secretion etc. Our results indicate that the short fiber protein can activate enterochromaffin cells to secrete serotonin, which may in turn cause the symptoms of gastrointestinal infection.

The results from these studies have helped in understanding and elucidating the biology of these adenoviruses and how they infect cells and cause disease.
Populärvetenskaplig sammanfattning

Adenovirus är virus som kan infektera olika organismer som sträcker sig från reptiler, fisk och fåglar till däggdjur. I denna avhandling fokuserar jag på adenovirus som smittar människor. Det finns flera typer av humana adenovirus som orsakar olika infektioner, såsom övre luftvägar, ögon eller maginfektioner. Mina studier fokuserar huvudsakligen på två humana adenovirus (typ 40 och 41) som orsakar mag- eller tarm-infektioner (gastrointestinal). De orsakar främst sjukdom hos spädbarn och är den tredje vanligaste orsaken till infektiösa diarréer (gastroenterit) i världen.


Humana adenovirus typ 40, 41 och 52 skiljer sig från andra humana adenovirus eftersom de har två olika fiberproteiner, en långt fiber och en kort fiber (de andra typerna har bara en fiber). Vidare har typerna 40 och 41 även en annorlunda pentonbas som troligen påverkar hur viruset tas upp av cellen. Den samma funktionen av det korta fiberproteinet och pentonbasen i den infektiösa cykeln hos de humana adenovirusen 40 och 41 är dock fortfarande en gåta. I denna studie undersökte vi funktionen av det korta fiberproteinet och pentonbasen
hos de humana adenovirusen 40 och 41. Vidare lyckades vi identifiera decellulära receptorerna för det långa och korta fiberproteinet från det nyligen isolerade humana adenoviruset typ 52.

I den första studien identifierade vi att integrinerna, $\alpha_6\beta_1$ och $\alpha_6\beta_4$, som interagerar med pentonbasen hos humana adenovirus 40 och 41. Integriner är proteinmolekyler på cellens yta som består av en $\alpha$-subenhet och en $\beta$-subenhet. Vi utförde bindnings- och infektionsexperiment på tarmceller och utvärderade interaktionen mellan pentonbasen och integrinerna. Liknande metoder användes sedan för att upptäcka en cellulär receptor för det korta fiberproteinet av typ 40 och 41, en proteinmolekylinnehållandes förlängda sockerkedjor och som kallas heparansulfat. Det långa fiberproteinet av humana adenovirus typ 52 interagerar med ett protein som kallades CAR (coxsackievirus och adenovirusreceptor) och det korta fiberproteinet interagerades med långa kedjor av sialinsyra. Dessa virus använder dessa identifierade receptorer för att binda till och komma in i cellen för infektion.

Därefter undersökte vi effekten av humana adenovirus typ 40 och 41 och deras kapsidproteiner på serotonin-utsöndrande celler som kallas enterokromaffin-celler i tunntarmen. Serotonin, en signalmolekyl, är involverad i flera gastrointestinala störningar och symptom som illamående, kräkningar, tarmutsöndring etc. Våra resultat indikerar att det korta fiberproteinet kan aktivera enterokromaffin-celler vilket leder till utsöndring av serotonin, vilket kan ge gastrointestinala infektions symptom. Resultaten från dessa studier har hjälpt till att förstå och belysa biologin hos dessa adenovirus och hur de smittar celler och orsakar sjukdom.
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AdPol</td>
<td>viral DNA polymerase</td>
</tr>
<tr>
<td>AdV</td>
<td>adenovirus</td>
</tr>
<tr>
<td>Bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BsDds</td>
<td>base dodecahedrons</td>
</tr>
<tr>
<td>CAM</td>
<td>cell adhesion molecule</td>
</tr>
<tr>
<td>CAR</td>
<td>coxsackievirus and adenovirus receptor</td>
</tr>
<tr>
<td>CBV</td>
<td>group B coxsackievirus</td>
</tr>
<tr>
<td>CHO</td>
<td>Chinese hamster ovary</td>
</tr>
<tr>
<td>CMV</td>
<td>cytomegalovirus</td>
</tr>
<tr>
<td>cryo EM</td>
<td>cryo electron microscopy</td>
</tr>
<tr>
<td>CSF</td>
<td>cerebrospinal fluid</td>
</tr>
<tr>
<td>DBP</td>
<td>DNA binding protein</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DPPC</td>
<td>dipalmitoyl phosphatidylcholine</td>
</tr>
<tr>
<td>dsDNA</td>
<td>double stranded deoxyribonucleic acid</td>
</tr>
<tr>
<td>DSG2</td>
<td>desmoglein 2</td>
</tr>
<tr>
<td>EC</td>
<td>enterochromaffin</td>
</tr>
<tr>
<td>ECM</td>
<td>extracellular matrix</td>
</tr>
<tr>
<td>EGC</td>
<td>enteric glial cell</td>
</tr>
<tr>
<td>EKC</td>
<td>epidemic keratoconjunctivitis</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>EM</td>
<td>electron microscopy</td>
</tr>
<tr>
<td>ENS</td>
<td>enteric nervous system</td>
</tr>
<tr>
<td>FIX</td>
<td>coagulation factor IX</td>
</tr>
<tr>
<td>FK</td>
<td>fiber knob</td>
</tr>
<tr>
<td>FX</td>
<td>coagulation factor X</td>
</tr>
<tr>
<td>GAG</td>
<td>glycosaminoglycans</td>
</tr>
<tr>
<td>Gal</td>
<td>galactose</td>
</tr>
<tr>
<td>GalNAc</td>
<td>N-acetylgalactosamine</td>
</tr>
<tr>
<td>GEMMA</td>
<td>gas-phase electrophoretic macromolecular analysis</td>
</tr>
<tr>
<td>GlcA</td>
<td>D-glucuronic acid</td>
</tr>
<tr>
<td>GlcN</td>
<td>glucosamine</td>
</tr>
<tr>
<td>GON</td>
<td>group of nine</td>
</tr>
<tr>
<td>GOS</td>
<td>group of six</td>
</tr>
<tr>
<td>HAdV</td>
<td>human adenovirus</td>
</tr>
<tr>
<td>HS</td>
<td>heparan sulfate</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HSPG</td>
<td>heparan sulfate proteoglycans</td>
</tr>
<tr>
<td>HA</td>
<td>hyaluronic acid</td>
</tr>
<tr>
<td>IdoA</td>
<td>L-iduronic acid</td>
</tr>
<tr>
<td>IGDD</td>
<td>isoleucine-glycine-aspartic acid-aspartic acid</td>
</tr>
<tr>
<td>KKTK</td>
<td>lysine-lysine-threonine-lysine</td>
</tr>
<tr>
<td>LDV</td>
<td>leucine-aspartic acid-valine</td>
</tr>
<tr>
<td>LF</td>
<td>long fiber</td>
</tr>
<tr>
<td>LFK</td>
<td>long fiber knob</td>
</tr>
<tr>
<td>MLP</td>
<td>major late promoter</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
</tr>
<tr>
<td>MTOC</td>
<td>microtubule organizing center</td>
</tr>
<tr>
<td>Neu5Ac</td>
<td>N-acetyl neuraminic acid</td>
</tr>
<tr>
<td>NPC</td>
<td>nuclear pore complex</td>
</tr>
<tr>
<td>PB</td>
<td>penton base</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PI3K</td>
<td>phosphoinositide-3-OH</td>
</tr>
<tr>
<td>pIIIA</td>
<td>protein IIIA</td>
</tr>
<tr>
<td>pIVa2</td>
<td>protein IVa2</td>
</tr>
<tr>
<td>pIX</td>
<td>protein IX</td>
</tr>
<tr>
<td>PPxY</td>
<td>proline-proline-x-tyrosine</td>
</tr>
<tr>
<td>pRB</td>
<td>retinoblastoma protein</td>
</tr>
<tr>
<td>PtDds</td>
<td>penton dodecahedrons</td>
</tr>
<tr>
<td>pV</td>
<td>protein V</td>
</tr>
<tr>
<td>pVI</td>
<td>protein VI</td>
</tr>
<tr>
<td>pVII</td>
<td>protein VII</td>
</tr>
<tr>
<td>pVIII</td>
<td>protein VIII</td>
</tr>
<tr>
<td>RC</td>
<td>replication competent</td>
</tr>
<tr>
<td>RCA</td>
<td>regulators of complement activation</td>
</tr>
<tr>
<td>RD</td>
<td>replication deficient</td>
</tr>
<tr>
<td>RGAD</td>
<td>arginine-glycine-alanine-aspartic acid</td>
</tr>
<tr>
<td>RGD</td>
<td>arginine-glycine-aspartic acid</td>
</tr>
<tr>
<td>RGN</td>
<td>arginine-glycine-asparagine</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>SCR</td>
<td>short consensus repeats</td>
</tr>
<tr>
<td>SF</td>
<td>shirt fiber</td>
</tr>
<tr>
<td>SFK</td>
<td>short fiber knob</td>
</tr>
<tr>
<td>Sia</td>
<td>sialic acid</td>
</tr>
<tr>
<td>SMP</td>
<td>submucosal plexus</td>
</tr>
<tr>
<td>SPR</td>
<td>surface plasmon resonance</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>ssDNA</td>
<td>single stranded deoxyribonucleic acid</td>
</tr>
<tr>
<td>STP</td>
<td>serine-threonine-proline</td>
</tr>
<tr>
<td>TP</td>
<td>terminal protein</td>
</tr>
<tr>
<td>TVD</td>
<td>threonine-valine-aspartic acid</td>
</tr>
<tr>
<td>VA</td>
<td>virus associated</td>
</tr>
<tr>
<td>WT</td>
<td>wild type</td>
</tr>
</tbody>
</table>
Aims of the thesis

Overall aim of this thesis is to explain the functions of capsid proteins of enteric human adenoviruses in their infectious cycle.

Aim 1: To investigate if enteric human adenovirus penton base proteins use integrins for entry and infection.

Aim 2: To identify receptors for the short fibers of human adenovirus types 40 and 41 and for the long and short fibers of type 52.

Aim 3: To study the mechanism whereby capsid proteins of enteric human adenoviruses cause clinical symptoms of gastrointestinal infection.
List of publications

I. Enteric species F human adenoviruses use laminin-binding integrins as co-receptors for infection of HT-29 cells.
   Manuscript.

II. Low pH primes short fibers of enteric human adenoviruses to use heparan sulfate as a cellular receptor.
    Rajan A*, Lenman A*, Trulsson F, Palm E, Mundigl S, Arnberg N.
    Manuscript.

III. Human adenovirus 52 uses sialic acid-containing glycoproteins and the coxsackie and adenovirus receptor for binding to target cells.

IV. Interaction of human enterochromaffin cells with human enteric adenovirus 41 leads to serotonin release and subsequent activation of enteric glia cells.

Not included in the thesis:

- Adenovirus-based vaccines for fighting infectious diseases and cancer: progress in the field.
  Majhen D, Calderon H, Chandra N, Fajardo CA, Rajan A, Alemany R, Custers J.
Human adenoviruses

Background

History

Human adenoviruses (HAdVs) were first isolated in 1953 from adenoid tissue derived cell cultures when Rowe et. al. (1) noted that these primary cell cultures underwent degeneration, which they recognized was caused by a transmissible agent. The discovery of these viruses in adenoid tissue was the basis for naming this virus family. Soon after, a HAdV was identified as the etiological agent of respiratory tract infections in military recruits (2). However, only a few types have been isolated from the adenoids/tonsils while most types target other cells, tissues or organs.

Adenoviruses (AdV) are widely used in basic research since they are small and have a simple genetic make-up, which allows for easy manipulation in vitro. HAdV research has helped in the better understanding of eukaryotic cell functions, including the discovery of messenger RNA (mRNA) splicing (3, 4) and the function of cell cycle regulating proteins such as p53 and the retinoblastoma protein (pRB) (5).

Taxonomy and Diversity

HAdVs belong to the family Adenoviridae and genus Mastadenovirus (6). Around 85 HAdV types have been identified so far and they have been grouped into seven different species, A to G (7). Earlier, HAdVs were classified based on serology, hemagglutination, oncogenicity in rodents, transformation of cultured primary cells and genome sequencing. Viruses that were resistant to neutralizing antibodies from one serotype were classified as a new serotype. However, nowadays, the different members of each species are referred to as ‘types’ and not ‘serotypes’ since classification is now based on genome DNA
sequences and bioinformatics analysis (8, 9). Hence, ‘types’ are defined by the amino acid sequence of the hexon capsid protein, which mainly contains the neutralizing epitopes. Since HAdVs are known to recombine among types and lead to new combinations of existing types (10), they are designated by the identity of the hexon gene (H), penton base gene (P) and fiber gene (F). For example, HAdV-85 is denoted as HAdV-H19/P37/F8, which is a recombinant virus with type 19 hexon, type 37 penton base and type 29 fiber.
Epidemiology and Clinical Features

AdVs are species specific, which means that HAdVs are not pathogenic to non-human mammals (as to the best of our knowledge) and animal AdVs are pathogenic only to their respective species. Even though Ads isolated from great apes are phylogenetically related to many HAdVs, interspecies infections are improbable (11). HAdVs account for 8% of clinically relevant viral diseases globally (12, 13). Several serologic surveys conducted in various parts of the world have established that species C HAdV types 1, 2 and 5 are the most seroprevalent (40-80%) in the human population (14-16). HAdV-6 (also species C), 11, 28, 35, 49, 50, 52 etc. are some types which have low seroprevalence (up to 30%) (17-26). The majority of the remaining HAdVs have a seroprevalence of 40 to 70%.

Symptoms of adenoviral infections in humans depend on the viral serotype, age and immune status of the individual. HAdVs commonly cause infections in the respiratory tract, eye and gastrointestinal tract. They can also infect the urinary tract and liver, although infrequently. HAdV-1, 2, 5 and 6 (all members of species C) cause about 7% of upper respiratory tract infections in children with symptoms such as nasal congestion, coryza and cough and 10% of lower respiratory tract infections causing pneumonia. HAdV-5 has also been speculated to cause whooping cough and pertussis-like symptoms. In addition to species C, species A, B1 (a subset of species B) and E cause acute respiratory disease in adults. (11).

HAdVs can cause infections of the eye as part of a respiratory-pharyngeal infection or separately (27). These infections are usually mild. A more contagious and severe form of follicular conjunctivitis is caused by HAdV-8, 19, 37, 53, 54 and 56 with edema of the eyelids, redness, pain, lacrimation and photophobia. This condition is known as epidemic keratoconjunctivitis (EKC) (28-38). EKC outbreaks most frequently originate from eye clinics, schools, military camps and other such crowded places (39).
Many HAdVs can replicate in the intestine and are excreted in stool but no correlation between clinical symptoms of gastroenteritis have been established with most HAdVs (40). Later, on examining stool smears of infected children by electron microscopy, two different enteric HAdVs – types 40 and 41 were associated with diarrhea (40-42). There have been several epidemiological studies on the enteric HAdVs over the years. The incidence of gastrointestinal infections caused by these viruses varies between locations with developed countries having a lower incidence and developing countries having a higher incidence of infections (43-51). They are, nonetheless, a major cause of infantile viral gastroenteritis worldwide after rotavirus and norovirus.

Members of species B2 (a second subset of species B) are associated with urinary tract infections and HAdV-11 (member of species B2) has also been reported to cause acute hemorrhagic cystitis in boys. Meningoencephalitis is rare, HAdV-3, 5, 6, 7, 12 and 26 have been demonstrated in the CSF. Other diseases that HAdVs have been isolated from are celiac disease, myocarditis and intussusception. In immunocompetent patients, the above-mentioned infections are usually mild and self-limiting. However, in neonates and immunosuppressed patients, HAdV infections can be fatal (52, 53).
Structural Proteins

HAdVs are non-enveloped double stranded DNA (dsDNA) viruses and are made up of three major and four minor capsid proteins arranged in an icosahedron (Figure 1). The inside of the capsid contains the DNA genome along with five core proteins and a viral protease (54). With a diameter of ~920 Å and molecular weight of ~150 megadaltons, HAdVs are one of the largest and most complex non-enveloped dsDNA viruses. The highest resolution structure of a HAdV particle (based on HAdV-5) is a 3.6 Å resolution cryo-electron microscopic (cryo EM) (54) and a 3.5 Å resolution X-ray crystallographic structure of the whole virion (55).

Figure 1: HAdV icosahedral shell organization and structure. Reprinted in accordance with the Creative Commons Attribution license (56).

Hexon (pII)

The hexon (Figure 2) is the most abundant capsid protein of the AdV particle. It is a trimer of three identical polypeptide chains and 240 such trimers make up the 20 faces of the icosahedral capsid (57). One
facet of the icosahedral capsid consists of 12 trimers of the hexon with nine hexon trimers (called Groups of Nine; GON) forming the central plate and the remaining three are adjacent to pentons from the adjoining three facets. Thus, there are five hexons, called peripentonal hexons, around each penton (called Group of Six together with the penton; GOS) (54).

Figure 2: Ribbon diagram of HAdV-5 hexon trimer with side view (upper panel and top view (lower panel). Hyper variable regions (HVRs) are highlighted (HVR1, red; HVR2, green; HVR3, pink; HVR4, light blue; HVR5, yellow; HVR6, blue; HVR7, cyan). Reprinted with permission from publisher (58).

Each hexon monomer has a dense pedestal base of two eight-stranded, antiparallel beta barrels (P1 and P2), which are stabilized by an internal loop shown by EM and X-ray crystallography. Three loops, namely, L1, L2 and L4 project away from the surface of the virion forming a tower domain (59). L2 and L4 loops coil to interact with the other two copies of the protein within the trimer making the structure of the trimer extremely stable. The pedestal regions are conserved among hexons of different HAdV species, however, variable regions
exist in L1 and L2 (60). Neutralizing antibodies are mainly generated against these variable regions and are HAdV-type specific. This epitope is historically designated as ε (epsilon) (61).

**Penton base (pIII)**

The penton base (PB) capsid protein (Figure 3) is present on each of the 12 vertices of the AdV particle. It is a pentameric protein with each monomer consisting of 470 to 570 amino acids (the size of the monomer varies with the different types of AdVs). The PB is highly conserved across all HAdV species having around 70% homology. Each monomer has two domains, a lower jellyroll-motif domain that faces the interior of the virus particle and an upper domain that consists of two loops – a variable loop (varies between AdV types) and an RGD loop (contains a conserved motif of the amino acids arginine (R)- glycine (G) - aspartic acid (D) and varies in sequence and length between types) (62). This RGD motif is present in all HAdV PBs except species F enteric HAdVs, which lack the RGD motif but instead have an arginine-glycine-alanine-aspartic acid (RGAD; in HAdV-40) and isoleucine-glycine-aspartic acid-aspartic acid (IGDD; in HAdV-41) motif. The RGD motif interacts with \( \alpha_v \) integrins, which are a co-receptor and facilitate internalization of the virus. It has been speculated that HAdV-40 and -41 do not bind integrins to enter host cells since they lack the RGD motif (63). As part of this thesis, we have studied co-receptor usage by HAdV-40 and -41 and identified \( \alpha_6 \) integrins as potential co-receptors and interacting partners for the PB of these viruses. The PB of AdV is also proposed to be involved in the release of the virus from endosomes (64-68).

Excess PBs that are produced during the viral infectious cycle can assemble to form dodecahedrons. Each dodecahedron is made from 12 pentameric PB proteins with or without the incorporated fibers. These protein complexes form highly regular structures. Particles that may include fibers are called penton dodecahedrons (PtDds) and the ones
without fibers are called base dodecahedrons (BsDds) (69). Some adenovirus serotypes such as HAdV-3 (member of species B) can spontaneously form dodecahedrons These structures have been proposed as alternative vectors to the whole AdV particle (70).

![Image of PB pentameric unit](image)

**Figure 3**: Surface representation (left) and ribbon diagram (right) of the side view of the PB pentameric unit. Each monomer is represented in a different color. Reprinted with permission from publisher (62).

**Fiber (pIV)**

The fiber protein sits on the PB at the 12 vertices of a HAdV particle. The fibers are trimeric proteins composed of identical monomeric polypeptide chains. The fibers have an N-terminal PB interacting domain, a long central shaft and a C-terminal head or knob domain (Figure 4) that is used for attachment to cells (71). The interaction between the five-fold symmetric PB and the three-fold symmetric fiber protein is due to (i) a hydrophobic ring around the central pore of the PB on its top surface, (ii) hydrophobic residues on the bottom of the fiber shaft and (iii) bendable N-terminal tails of the fiber. Neighbouring subunits of the PB form grooves into which the N-terminal tails insert (62, 72). The shaft of the fiber protein is composed of many copies of a ~15-residue structural motif (73).
Difference in the number of repeats of this structural motif causes variation between the shaft lengths of different types of HAdVs. For example, HAdV-12 has a shaft with 23 repeats whereas HAdV-3 only has six. Roelvink and colleagues claim that shaft length is responsible for the divergent tropism of the different types (74). Along with shaft length, flexibility of the fiber protein is also proposed to contribute to receptor interactions on the cell surface (75).

HAdVs use the C-terminal fiber knob (FK) to attach to cell surface receptors. Most HAdVs have one fiber protein. However, HAdV-40, -41 and -52 have two fibers with different lengths and they are designated as long and short fibers (LF and SF) (76-78). For these HAdVs, each PB at the vertex incorporates only one of the two fiber proteins. For HAdV-52, there are equal number of long and short fibers (79) but for HAdV-40 and -41, there are more short fibers than long fibers (80).

Figure 4: Space-filling model of HAdV-5 fiber protein showing knob, shaft and tail domains. Each monomer is represented in a different colour. Reprinted with permission from publisher (72).
**Minor Capsid Proteins**

**Protein VI (pVI)**

Protein VI (pVI) is present on the inside of the HAdV capsid. It binds to a loop in the inner cavity of the hexon trimer and to the dsDNA genome, thus bridging the core to the icosahedral shell (54, 81, 82). Each virus particle contains 60 hexamers of pVI (54). Protein VI is a multifunctional protein with roles in many stages of the viral replication cycle. During entry, the N-terminal of pVI alters the endosomal membrane causing the virus to be released into the cytosol (83-85). It contains a conserved PPxY motif that interacts with ubiquitin ligase and facilitates trafficking along microtubules to the nucleus (86). A nuclear localization signal present on its C-terminal directs newly synthesized hexon to the nucleus (81).

**Protein IIIA (pIIIA)**

Protein IIIA (pIIIA) is a minor capsid protein located on the inside of the capsid arranged in a ring under each vertex formed by five pIIIA monomers. The N-terminal domain secures the interaction between peripentonal hexons and PB (87). Two functions have been proposed for pIIIA: stabilizing the vertex region and packaged genome during assembly of viral particles and signaling for release of the vertex and genome during uncoating (88, 89).

**Protein VIII (pVIII)**

The function of pVIII is not well characterized. Squeezed between pIIIA and hexons, the 120 copies of pVIII stabilize the hexons together with pIX. (55).

**Protein IX (pIX)**

Among all the genera present under *Adenoviridae*, pIX is only present in *Mastadenovirus*. Each particle contains 240 copies, which are located on the outside of the capsid (54). A pIX monomer is made up
of 140 amino acid residues and consists of an N-terminal domain, a rope domain and a helix bundle. Protein IX holds the GON hexons together to form stable substructures that do not separate upon mild dissociation conditions (90). Although pIX is needed for capsid stabilization, pIX-deleted mutants can assemble into viral particles, however, they have lower thermo stability (91). It has been observed in structural studies that at least one of pIX’s domains is exposed on the capsid, which has made it a prospective candidate for modification of vectors for retargeting, immunization etc (92).

**Core Proteins**

The core of an adenoviral particle consists of the viral DNA genome; proteins V, VII, μ, IVa2, terminal protein (TP) and a viral protease (Figure 5). Protein VII, the most abundant core protein, along with pV and μ are arginine-rich proteins with a basic charge, which interacts with the DNA and compresses it within the core (93-97). Protein V also interacts with pVI, the DNA (98) and to the N-terminus of each PB monomer, which tethers the protein capsid to the DNA-protein core (99). Each viral particle has less than ten copies of pIVa2, which is present at one vertex of the icosahedron. Protein IVa2 is required for packaging viral DNA into the capsid during virion assembly by binding to a packaging signal on the genome (100). The TP acts as a primer for DNA replication and the two copies per viral particle are anchored to the 5’ ends of the genome. There are also around 10 molecules of the p23 viral cysteine protease present in the core of each viral particle and it functions during assembly/maturation and during disassembly/egress of the virions from endosomes. The viral cysteine protease cleaves virion protein precursors before assembly into a mature viral particle (101).
Figure 5: Schematic figure of HAdV core proteins. Reprinted in accordance with the Creative Commons Attribution license (56).

Table 1: Summary of HAdV-5 structural proteins and copy number per virus particle. Adapted from (54).
Infectious Cycle

The infectious cycle of HAdVs begins with a two-step cell entry pathway. The first step involves attachment of the fiber to a primary receptor on the cell surface, which is followed by the second step where interaction of the PB with cellular integrins causes endocytosis of the viral particle. After incomplete dissociation and release from endosomes, the partially disassembled virion is transported to the nucleus, where it releases its viral DNA. This DNA is replicated in the nucleus. Capsid proteins which are synthesized in the cytoplasm are imported into the nucleus, where it encapsidates the newly replicated DNA. Finally, progeny virions are released upon lysis of the cell. In this section, steps from attachment to release will be discussed in more detail, with a focus on the initial steps. The entire replication cycle takes approximately 24 to 36 hours and each infected cell can produce up to $10^5$ new virions (7).

**Attachment and Attachment Receptors**

![Figure 6: HAdV attachment receptors. Modified and reprinted with permission from the publisher (102).]
Attachment of HAdVs to the host cell is the first step of the infectious cycle. The fiber protein is the first viral component to interact with the attachment receptor on the cell surface (Figure 6). HAdVs can bind to protein receptors – membrane-bound or soluble – and glycan receptors depending on the type and species of HAdV. The different attachment receptors are described below.

**The coxsackievirus and adenovirus receptor**

The coxsackievirus and adenovirus receptor (CAR) is a membrane-bound protein comprising of a long cytoplasmic tail, a transmembrane region and an extracellular region containing two immunoglobulin (Ig)-like domains (103, 104), denoted as D1 and D2. The 46 kDa-CAR protein is localized along the basolateral surface of tight junctions, which are cell adhesion contacts between adjacent cells (105, 106). CAR plays a role in the barrier function of tight junctions by controlling paracellular transport of water, ions and other solutes by forming homodimers with CAR molecules on adjoining cells. CAR is expressed in large quantities in the heart, pancreas, the central and peripheral nervous systems, prostate, testis, lung, liver and intestine, but it is absent or expressed sparsely on hematopoietic cells and adult muscle (107).

HAdVs and group B coxsackieviruses (CBV) share CAR as their attachment receptor (108). CAR was initially discovered as a receptor for HAdV-2 and -5 in 1997 (104). Later, CAR was also shown to be the attachment receptor for HAdV species A, D, E, F and G (74, 79). HAdVs and CBV bind to the D1 domain of CAR albeit to different epitopes (109, 110). CAR acts merely as an attachment receptor since it has been shown that neither the cytoplasmic part nor the transmembrane region is needed for HAdV infection (111). CAR binds to highly conserved (among the CAR-utilizing HAdV types) amino acid residues comprising exposed loops on lateral surfaces of the fiber knob domain (74, 110, 112, 113). The usage of CAR as an attachment receptor depends on the length and flexibility of the binding fiber. The number of repeating structural motifs determines
the length of the shaft domain within which there are distinct segments that attribute to the flexibility of the fiber shaft (71, 75). Altering the number of fiber shaft repeats from the normal 21.5, reduces CAR binding (75, 114, 115). The affinity of CAR-fiber knob interaction is high with $K_D \approx 1 \text{nM}$ (116).

CAR is commonly found in tight junctions and also expressed mostly along the basolateral surface of epithelial cells. This location of CAR is not ideal for attachment by a virus approaching cells from the apical side. However, an infecting virus can potentially access and attach to CAR if there are lesions in the epithelium (117). Also, a less common isoform of CAR (CAR$^{Ex8}$) can be exposed on the apical side of airway epithelial cells and can support viral attachment (118). One of the chief functions of cell adhesion molecules (CAMs) such as CAR is preserving tissue integrity and promoting migration of immune cells, which it achieves, for example, by modulating the cell surface expression of CAMs through endocytosis (119). Diverse viral families exploit this function of CAMs, as this could be a strategy to gain entry into the cell. Finally, it has been suggested that CAR has an additional function, where it can promote release of progeny virions by hindering CAR dimerization at tight junctions. The excess fiber proteins synthesized during viral replication that are not incorporated into virions are free to bind CAR monomers at the tight junctions inhibiting CAR-CAR dimerization and increasing intercellular space leading to dissemination of newly assembled virions (117).

**CD46**

CD46 or membrane co-factor protein is a membrane-bound receptor, which mediates cleavage of complement components C3b and C4b by factor I thereby inactivating them and rendering protection to the healthy host cell against damage by complement. CD46 is a member of a family of proteins called regulators of complement activation (RCA) and is expressed on all nucleated cells (120, 121). CD46 has an extracellular domain, a transmembrane region and a cytoplasmic domain. The extracellular domain is made up of four short consensus
repeats (SCRs) that are arranged in an elongated fashion, three of which are N-glycosylated. The region closest to membrane is rich in serine-threonine-proline (STP) residues and is named the STP region. The STP region is heavily O-glycosylated. Most human tissues express four major isoforms of CD46 – BC1, BC2, C1 or C2. These isoforms differ in their cytoplasmic tail (C1 or C2) or in their STP extracellular region (BC1 or BC2; with or without B segment) (122).

Unlike the other HAdV species, species B HAdVs use CD46 as their primary attachment receptor (123, 124). HAdV types 16, 21, 35 and 50 exclusively use CD46 as a receptor while HAdV-3, -7 and -11 can use CD46 in addition to other receptors (123-126) and the interaction can be of high or low affinity. HAdV-7 and -14 bind CD46 but the affinity of the interactions makes CD46 an unlikely receptor for productive infection (127). HAdV-37 is one of several species D HAdVs that has also been shown to bind CD46 (128). For high affinity interactions, it is known that HAdVs bind to the N-terminal domains (SCR1 and SCR2) of CD46 with two of the three fiber knob monomers. The angle between SCR1 and SCR2 changes after fiber knob binding, which, in theory, may allow the third fiber knob monomer to engage in binding thus increasing avidity (129-131). The SCR domains of CD46 serves as a receptor for a variety of pathogens such as human herpes virus 6 (SCR2 and 3), laboratory strains of measles virus (SCR1 and 2), Streptococcus pyogenes (SCR3 and 4) and Neisseria gonorrhea (SCR3 and STP domain) (132).

Desmoglein 2

Desmoglein 2 (DSG2) is a calcium-binding transmembrane glycoprotein of the desmosome junctional complex and belongs to the cadherin family of proteins. Within the desmosome, DSG2 helps in tethering cellular intermediate filaments to the plasma membrane and in forming strong cell-cell adherent junctions (133). DSG2 is expressed on most human tissues and is used as a cellular attachment receptor for species B HAdV-3, -7, -11 and -14. HAdV-3 interacts with DSG2 through its PB proteins (shown by using PtDds) with an
affinity of 2.5 nM (134). HAdV-11 uses DSG2 as a cellular receptor, in the absence of CD46. Similar to CAR, DSG2 is localized in cell junctions of epithelial cells, which makes it inaccessible for an incoming virus to attach to the host cell. In similarity with CAR-binding fibers, dodecahedrons can also be secreted from infected cells and can bind DSG2 in the desmosome complex, which causes intracellular signaling and degradation of the extracellular domain of DSG2 thus impeding the formation of junctions with DSG2 on neighboring cells and allowing spread of viral particles (134-136).

**Sialic acid-containing glycans**

Sialic acids (Sia:s) are monosaccharides that belong to a family of nine-carbon acidic sugars. Within this family, N-acetyl neuraminic acid (Neu5Ac) is the most predominant among higher animals. Sia:s are usually present at the outermost ends of N-glycans, O-glycans and glycosphingolipids where they can be modified in several different ways such as by addition of acetyl, hydroxyl, lactyl, methyl, sulfate, phosphate and many other groups (137) giving rise to over 50 variants (138). Further variation in Sia:s originates from different α-linkages between its carbon at position 2 to the carbon at position 3 or 6 of galactose (Gal) or to the carbon at position 6 of N-acetylgalactosamine (GalNAc) residues. Homopolymers of Sia:s where monosaccharides link to carbon at position 8 form polysialic acids, which are rare in animal glycoproteins (139, 140). N-acetyl, carboxylate, glycerol and hydroxyl groups protrude from Neu5Ac, which permits a large number of hydrogen bonding, salt-bridge formation and non-polar interactions. The possibility of numerous interactions, along with the location and accessibility of Sia at the terminus of a glycan makes it an appealing target for pathogens. Some examples of pathogens using Sia as receptors are HAdVs, coxsackievirus A24 variant, influenza virus A, B and C, JC and BK polyomaviruses, rotaviruses, *Helicobacter pylori*, *Plasmodium falciparum* etc (137, 141). The negative charge at the cell membrane generated by terminal Sia residues on glycans is responsible for many biological functions such
as causing a repulsion between cells (142, 143), shielding cell surface proteins from degradation caused by proteases (144) and using Sia patterns to differentiate between self and non-self structures (140).

Epidemic keratoconjunctivitis causing HAdV types – 8, 37 and 64 mainly use Sia as a receptor instead of CAR (145). However, HAdV-37 fiber knob domain is capable of binding CAR but does not use it as a functional receptor due to the short (7.5 β-repeats) and inflexible shaft (75, 113). The interaction of the fiber knob and Sia is of low affinity (mM). An exception is the affinity of HAdV-37 FK with GD1a, which is much lower, i.e., in μM levels (146). For HAdV-37, Sia binds to the pocket formed between the three fiber monomers (147, 148).

**Heparan sulfate proteoglycans**

All cells express heparan sulfate (HS), which is a type of glycosaminoglycan (GAG) made up of alternating uronic acid (D-glucuronic acid [GlcA] or L-iduronic acid [IdoA] and D-glucosamine (GlcN) units (149). These GAG chains are present as covalent complexes with specific proteins called core proteins. HS-containing chains with the complexed core protein make up the superfamily of HS proteoglycans (HSPG). Proteoglycans can be membrane bound or a constituent of the extracellular matrix (ECM). Syndecans (syndecan-1, -2, -3 and -4) (150, 151) and glypican (1 to 6) (152) are two plasma membrane bound core proteins that contain two to five HS chains bound to the extracellular ectodomain. Perlecan, epican, agrin, betaglycan, collagen XVIII are examples of other proteoglycans (149, 153, 154). HSPGs play fundamental roles in a multitude of organisms (e.g. Drosophila melanogaster, Caenorhabditis elegans, Hydra etc.). Diversity in HS chains can be due to variably sulfated or non-sulfated GlcA/IdoA and GlcN residues – disaccharide GlcA-GlcNS3S occurs mostly in endothelial cells and connective tissue mast cells (155, 156) while IdoA2S-GlcNS3S is predominantly expressed in the kidney (157).
HSPGs participate in myriad functions in the cell (Figure 7). Perlecan, agrin, collagen XVIII define the structure of the cellular basement membrane and provide a matrix for cell migration by interacting with other matrix components. HSPGs are present in secretory vesicles where they regulate many biological activities after secretion, package granular contents and maintain proteases in an active state. HSPGs prevent degradation of cytokines, chemokines and growth factors by binding to them. Membrane-bound proteoglycans play an essential role in cell:ECM attachment, cell:cell interactions and adhesions, and cell motility by collaborating with integrins. The above mentioned functions are some examples of the wide range of cellular functions of HSPGs (158).

Many pathogens have been shown to interact with HSPGs such as herpes simplex virus, dengue virus, cytomegalovirus and Chlamydia trachomatis (159-163). It has been proposed that HAdV-2 and -5 use HSPG as a receptor. HAdV-HSPG interaction involves a basic KKTK motif present in the third β-repeat of the fiber shaft although direct binding has not been confirmed (164). HSPGs have also been proposed as a co-receptor for HAdV-3 and -35 (165).

**Attachment via soluble components**

Several non-fiber mediated attachment and entry mechanisms affect HAdV infection in vivo. Coagulation factor X (FX) binds to the hexon subunit of most HAdVs, which permits direct binding to HSPG on liver hepatocytes in animal models (166). Coagulation factor IX (FIX) binds to HAdV-31 hexon with a high affinity of 3 nM (167). Blood components such as lactoferrin (168), complement components C3 and C4Bp (169) also amplify hepatocyte transduction. Dipalmitoyl phosphatidylcholine (DPPC) constitutes the pool of phospholipid surfactants necessary for lung function and binds to HAdV-5 hexon protein leading to entry of the virus independently of cellular receptors (170-173).
Internalization and Entry Receptors

Integrins

Integrins are a large family of transmembrane glycoproteins that are produced by a wide range of organisms from sponges to mammals (175). Integrins are essential cellular components and they regulate several functions such as cell adhesion, migration and signaling. Integrins carry out these functions by interacting with various ligands such as ECM glycoproteins, cellular receptors and cell adhesion molecules (176-178). Integrins are heterodimeric receptor molecules composed of an α and a β subunit (Figure 8). So far, 18 α and 8 β subunits have been identified that can dimerize and form at least 24 different integrin dimers. Each integrin subunit consists of an extracellular, a transmembrane and a cytoplasmic domain. The extracellular domain is the largest part while the cytoplasmic domain
is short (10-70 amino acids) except for the cytoplasmic tail of β4, which is over 1000 amino acids long (179). Signaling by integrins are accomplished in two ways: inside-out signaling, where alteration in the cytoplasmic tails transmits signals to the extracellular domain or by outside-in signaling, where activation of the integrin by ligand binding triggers transmission of signals from outside to inside (180).

Many pathogens interact with integrins to penetrate the cell membrane for successful infection. Herpes simplex virus 1 (181), coxsackievirus A9 (182, 183), rotavirus (184-187) and ebola virus (188) are some examples. HAdV interaction with integrins is required for internalization of the virus into the host cell. The PB of HAdVs interacts with integrins (K_D ≈ 50-80 nM) through the RGD motif on one of its exposed loops (62, 189-193). This interaction is facilitated by the bending of the flexible fiber (75). HAdVs with short fiber shafts may interact with integrins without bending since the PB will be closer to the cell surface than viruses with long shafts (194). Four to five integrins can be bound per PB molecule (RGD loops of each monomer are ~60 Å apart). Since, this interaction is needed in addition to attachment to the primary receptor, integrins are known as co-receptors for HAdVs. Several integrins have been identified as co-receptors or secondary receptors such as αvβ1, αvβ3, αvβ5, α3β1, and, α5β1 (193, 195-198). Another function of integrins in the HAdV infectious cycle is that of endosomal release (192, 199). This has been shown for αvβ5 but not for αvβ3 since β3 lacks a TVD motif required for this mechanism (200, 201).
Once integrins bind to a pentameric PB, they cluster and activate phosphoinositide-3-OH kinase (PI3K), p130CAS and Rho family GTPases (202). This causes actin polymerization, which prepares the cell for endocytosis via clathrin-coated pits (107). Species B HAdVs internalize through another mechanism of endocytosis known as macropinocytosis (203, 204). The virus uncoats at the cell surface, fibers are lost and endocytosis of a fiber-less particle occurs (205). After endocytosis, clathrin-coated vesicles mature into endosomes in the cell cytoplasm (206). Structural studies (55, 62, 72) have shown that expansion of the central pore of the PB enables the fiber’s release. The acidic environment of the endosomes facilitates the release of PB, peripentoneal hexon, pIII A, pVI and pVIII (207). Acidification also causes activation of the viral protease and cleavage of pVI (208). Protein VI cleavage in turn mediates disruption of the endosomal membrane and the partially uncoated virions is delivered into the

Figure 8: Integrin structure showing bent (left) and upright (right) conformations. Reprinted with permission from the publisher (180).
cytosol (83). Next, these particles travel along microtubules to the nucleus (209-211). From the microtubules, the subvirion particles move to the microtubule-organizing center (MTOC) and then associate with nuclear pore complexes (NPCs) (206). This association is due to an interaction between hexon trimers and the NPC cytoplasmic filament protein CAN/Nup214 (212). Further uncoating of the particles occurs at the NPC to allow the DNA to pass into the nucleus (206). Hexon and pIX remain at the NPC while pVII bound viral DNA enters the nucleus (206, 213-215). The steps from attachment to nuclear entry of DNA are illustrated in figure 9.

**Genome Transcription and Replication**

AdVs have a linear, dsDNA genome and the length between the different types varies from 26 kb to 45 kb (216). The replication of adenoviral genomes includes characteristic features such as accurate temporal control of viral gene expression and a protein priming mechanism of viral gene expression. HAdV genomes have five early

![Diagram of cellular internalization and trafficking of HAdVs.](image)

**Figure 9**: Cellular internalization and trafficking of HAdVs. Reprinted with permission from the publisher (217).
transcription units (E1A, E1B, E2, E3 and E4), four intermediate transcription units (IX, IVa2, L4 intermediate and E2 late) and a late transcription unit that generates five late mRNAs – L1 to L5, which is generated as one transcript from the major late promoter (MLP) (Figure 10). Three viral proteins – TP, viral DNA polymerase (AdPol) and viral single-stranded DNA (ssDNA) binding protein (DBP), are requisite for viral DNA replication. DNA replication originates at inverted terminal repeat sequences, which can range in size from 36 to 200 bp between the genomes of the different AdVs (216). Single strands of viral DNA that are displaced during asymmetric strand synthesis, circularize by base pairing of their terminal repeat sequences and function as replication origins. Genome transcription and replication occurs as follows (218):

- Host cell’s RNA polymerase II transcribes the E1A gene immediately following infection. These mRNAs are alternatively spliced and exported to the cytoplasm where E1A proteins are produced by the host translation machinery.

- The E1A proteins are imported into the nucleus and activate transcription of E1A, E1B, E2 early, E3 and E4 regions. Also, during this early phase, host RNA polymerase III transcribes viral-associated (VA) genes encoded by E2. These pre-mRNAs are translated to generate proteins essential for viral replication (219, 220).

- The translated viral replication proteins are imported into the nucleus. The pre-TP preserves the integrity of the viral chromosome’s terminal sequence and functions as a primer to initiate DNA synthesis (221-224). The AdPol is a 140-kD protein having both 5´ to 3´ DNA polymerase activity and 3´ to 5´ proofreading, exonuclease activity (225). The DBP tightly binds and stabilizes ssDNA in a sequence-independent manner and helps unwind and separate the two DNA strands (226). Two host-cell transcription factors – NF1 and OCT1 bind to the left end of the terminal repeat on the viral DNA, enhancing efficiency of replication initiation.
- After replication initiation, chain elongation continues with the help of AdPol, DBP and a cellular topoisomerase (227). Topoisomerase I increases the synthesis speed of new chains and can be needed to solve a DNA structural problem after extensive replication. Newly replicated viral DNA molecules serve as templates for subsequent rounds of replication and for transcription of late genes.
- Replication of viral DNA activates expression of proteins IX and IVa2. L1 to L5 mRNAs are also transcribed during this time from the MLP.
- L4-100 kDa protein escorts hexons for assembly, which along with the other structural proteins is imported into the nucleus for assembly and viral DNA packaging.

**Figure 10**: HAdV-5 genome structure and organization. Reprinted with permission from the publisher (218).

**Assembly and Release**

Viral DNA replication and production of vast amounts of structural proteins triggers the suppression of host-cell mRNA translation and enhances translation of late viral mRNAs. The newly synthesized capsid proteins are transported to the nucleus, where adenoviral
particles are assembled. As mentioned earlier, the L4-100 kDa protein (encoded by adenoviral late genes under the MLP) helps in the proper folding of hexons and functions as a starting point for the assembly of hexon trimers (228). Pentameric PB proteins and trimeric fibers join together to form penton capsomeres (229, 230).

The HAdV genomes comprise of many repeats of a cis-acting sequence between the left terminal repeat and the first coding sequence for the E1A gene. This sequence (5´-TTTG-N₈-CG-3´) is needed for correct packaging of viral DNA into infectious virions and is located within several hundred base-pairs of one end of the chromosome to ensure proper packaging (231). Viral proteins IVa2, L4-22 kDa, and L1-52/55 kDa bind the packaging sequence and facilitate viral DNA packaging into pro-capsids (232-236). Protein IIIa is involved in the packaging process due to its interaction with L1-52/55 kDa (237). Protein VII, which is the major histone-like core protein, interacts with viral DNA in the late phase, along with IVa2 and L1-52/55 kDa (38, 366, 393). Viral DNA (along with the attached pre-TP is packaged) complexed with pVIII, pV and precursor of μ is packaged. The p23 viral cysteine protease cleaves protein VI, VII, VIII, μ and TP precursors thus completing the assembly of mature, infectious virions.

Adenoviruses are lytic viruses and to release progeny virions, the infected cells need to lyse. This takes place through several processes such as: (i) cleavage of cellular cytokeratin K18 by viral protease L3-23K (238), (ii) cell death by the E3 11.6 kDa protein (also known as the AdV death protein), where it interacts with the cellular protein MAD2B which has a regulatory function in the cell cycle (239, 240), (iii) tight-junctions between cells are disrupted when protein E4orf1 interacts with PDZ-domain-containing members of the membrane-associated guanylate kinase protein family leading to progeny virus dissemination (241, 242). The HAdV-2 infectious cycle from attachment to assembly and release is illustrated in figure 11.
Figure 11: HAdV-2 infectious cycle from attachment to assembly and release. Reprinted with permission from the publisher (218).
Enteric HAdVs: dissimilarities and challenges

Enteric HAdVs differ in their structure, cellular uptake mechanism and growth characteristics in cell culture compared to the other HAdVs especially the most well characterized ones – HAdV-2 and HAdV-5. As described earlier, enteric HAdVs are structurally different since they have two types of fibers – a long fiber and a short fiber (76, 77) and the RGD motif required for interaction with αv integrins is absent on their PB protein (243). Enteric HAdVs mainly cause disease in the intestinal tract but not in other organs in the body (43). It has been shown in A549 cells (an epithelial human lung carcinoma cell line) that these viruses have a delayed uptake mechanism (243), which affects replication and spread of the viruses to other cells. Two studies (243, 244) have shown that the usual adsorption time of one hour is insufficient for productive infection of A549 cells and that for enteric HAdVs, the optimal time for internalization and entry is ≥4 hours. The E1 proteins of HAdV-40 and -41 are poorly expressed, which hinders their replication cycle (245-247). HEK293 cells express E1 proteins of HAdV-5 and hence are the preferred cell line for enteric HAdV propagation. Although even in these cells, the viruses display fastidious growth patterns (248) and have up to $10^4$ times higher particle to infectivity ratio as compared to HAdV-5 (63). Brown et. al. (63) compared amounts of cell-associated and released virions in HeLa and HEK293 cells for HAdV-5, -40 and -41. Fewer infectious virions were detected in the culture fluid for the enteric HAdVs in HeLa cells than HEK293 cells and virus yield for the enteric HAdVs was $10^5$ times lower than HAdV-5, which indicated an impediment in efficient release of HAdV-40 and -41 virions.

The fastidiousness, orders of magnitudes higher particle to infectivity ratios, blocks in virus entry and release, lack of appropriate in vivo models makes studies on the enteric HAdVs more challenging than the other human HAdVs.
Clinical applications

The chief clinical application of HAdVs is as a viral vector. AdV vectors are the most frequently used vectors for cancer therapy. They are also widely used in the development of gene therapies, where the vector expresses a non-mutant protein to correct a genetic defect and as vaccine vectors, where they express a foreign antigenic protein. Counting past and current clinical trials, HAdV vectors have been used in more than 400 trials (11, 249).

There are two kinds of HAdV vectors based on HAdV-5 – replication-defective (RD) and replication-competent (RC). RD vectors can have some or all of their early transcript regions deleted (250, 251) or retain only the DNA replication origins and ~500 bp of the left end of the genome containing the packaging sequence while the remaining genome is deleted (252). RC vectors or oncolytic vectors work on the principle that cancer cells are more permissive to HAdV replication as compared to non-cancerous cells. Thus, RC vectors lyse cancer cells at the end of their infectious cycle (253). ONYX-015 is an RC vector in which the gene encoding E1B-55K is deleted. These vectors cannot replicate in normal cells but they can replicate in many cancerous cells that are able to compensate for the function of E1B-55K of exporting late mRNAs from the nucleus into the cytoplasm (254, 255). Another RC vector called Oncorine was approved for use in China (256).

HAdV-5 based vectors, although, have some shortcomings (166, 257-260). Having high seroprevalence worldwide, accumulating in the liver after binding to coagulation factors on intravascular administration, eliciting a robust immune response and poor accessibility of CAR are some drawbacks. These issues have led to strategies such as shielding capsid proteins using polyethylene glycol (PEG) to reduce recognition by neutralizing antibodies (261), replacing the hexon with hexon another HAdV type (e.g. HAdV-3) (262) to minimise liver accumulation and developing vectors based on rare human types or non-human types (263).
Enteric Nervous System

The enteric nervous system (ENS) is a complex network composed of neurons and ganglia (Figure 12). This neural network is embedded in the walls of the whole gastrointestinal tract and its associated glands such as the salivary glands and pancreas. The ENS plays a role in every aspect of gut physiology and pathophysiology. It is a component of the autonomous nervous system and regulates intestinal motility, secretions, blood flow and immune responses. The neurons and enteric glial cells (EGCs) form plexuses throughout the length of the intestine. Two significant plexuses are the myenteric plexus and the submucosal plexus. The myenteric plexus is present throughout the digestive tract – from the esophagus to the rectum – between the outer longitudinal and circular muscle layers. The submucosal plexus is mainly in the intestines and is constituted of two layers. One layer is present on the non-luminal (abluminal) side of the muscularis mucosa while the second layer is adjacent to the luminal side of the circular muscle coat. The myenteric plexus has a higher density of neurons than the submucosal plexus (264).

We entirely depend on the gastrointestinal system to breakdown food, absorb nutrients and excrete waste. This system also needs to shield itself from toxins, physical damage and commensal bacteria. These complex and indispensable processes need substantial neuronal coordination, which makes the ENS a key component of a normal functioning intestine. The ENS has numerous functions such as regulating gastric acid secretion, controlling fluid movement over the intestinal epithelial lining and interacting with endocrine and immune systems of the gut (265).

Enterochromaffin (EC) cells are enteroendocrine cells in the small intestine and their primary function is to secrete serotonin for modulating gastrointestinal neurons. They activate the ENS through enteric glial cells and further via the extrinsic vagal afferents to the brain by producing and secreting serotonin through the basolateral
surface of the intestinal mucosa (266-268). Another main function of EC cells is to regulate water and electrolyte secretion through the intestinal epithelium. This secretion through the luminal side of the epithelium is under the control of the vagus nerve (269). Serotonin secretion is involved in modulating intestinal fluid motility and secretion, nausea, vomiting and gastroenteritis. Serotonin also plays a role in rotavirus pathology (270-272). It has been shown that the non-structural NSP4 protein of rotavirus can stimulate release of serotonin from EC cells, which is a plausible mechanism for rotavirus induced diarrhea and vomiting (270). Thus, the ENS and cells activating the ENS are important and influential players in gut physiology and viral gastroenteritis.

Figure 12: The enteric nervous system with myenteric plexus and submucosal plexus (SMP) formed by neurons and enteric glial cells within the intestinal tract wall. Reprinted with permission from the publisher (265).
Results and Discussion

Paper I

Enteric species F human adenoviruses use laminin-binding integrins as co-receptors for infection of HT-29 cells.


Manuscript.

The purpose of this study was to elucidate the role of the PB protein of enteric HAdV types 40 and 41. The PBs of the majority of HAdVs bind integrins on the cell surface and mediate the virus’ entry into the host cell and its release from endosomes. Cell:cell and cell:ECM adhesion are some of the important functions of integrins. ECM proteins such as vitronectin, fibronectin, laminin, collagen etc. are integrin ligands and they can bind integrins via conserved motifs such as RGD. Many viruses, including HAdVs, have evolved to mimic ECM proteins and encode similar or identical motifs. Some integrins that HAdVs bind to are $\alpha_\text{v}\beta_3$, $\alpha_\text{v}\beta_5$, $\alpha_\text{v}\beta_1$, $\alpha_5\beta_1$ and $\alpha_3\beta_1$. The PB of all HAdVs except types 40 and 41 contain an RGD motif. It was interesting to investigate the function of enteric HAdV PBs because they lack an RGD motif and there were no studies describing the interacting partners or function of these proteins. In this paper, we first evaluated transduction of a panel of Chinese hamster ovary (CHO) cells using a HAdV-41 GFP expressing vector. These CHO cells over-express alpha subunits of human integrins that bind ECM proteins. We validated the expression of the corresponding integrins on the different CHO cell lines using monoclonal antibodies against the alpha subunit of the integrin by flow cytometry. With the HAdV-41 GFP vector, we noted enhanced transduction on cells over-expressing $\alpha_2$, $\alpha_3$, $\alpha_6$, $\alpha_5$, $\alpha_8$ and $\alpha_9$ integrins as compared to the parental cell lines that do not express human integrins. $\alpha_2$-, $\alpha_3$- and $\alpha_6$-containing integrins belong to the group of laminin-binding integrins. $\alpha_5$- and $\alpha_8$-containing integrins belong to the group of RGD-binding integrins and
α9 belongs to the LDV-binding group of integrins, where LDV stands for the amino acids, leucine-aspartic acid-valine. From these results, we could interpret that the above-mentioned integrins may possibly be involved in the early stage of enteric HAdV infectious cycle. Increased transduction of α5- and α8-expressing CHO cells was surprising. But it is also known that RGD binding integrins are promiscuous in their ligand usage and they are known to bind to a wide variety of ECM ligands. Further, we mainly focused on the group of laminin-binding integrins since only enteric HAdVs possess ASK/IEQ (HAdV-40) and ASIQK/IEK (HAdV-41) motifs on their PB. These motifs are present in the PB loop corresponding to the RGD loop for other HAdVs and are similar to the SIKVAV/IEK/LEQ motifs on laminin chains (273, 274). To analyze if laminin-binding integrins are required in the entry and infection of enteric HAdVs, we used recombinant soluble laminin 332 (ligand for integrins α3β1 and α6β4) and laminin 511 (ligand for integrins α2β1, α3β1 and α6β4) to block cellular integrins on HT-29 cells. HT-29 cells are a human colon carcinoma cell line expressing small intestinal markers (275). We pre-incubated these cells with the soluble laminins before analyzing infection and internalization and we observed a reduction in both infection and internalization in cells that were pre-treated with soluble laminins as compared to untreated cells. Prior to infection/internalization assays we ensured that HT-29 cells express the candidate laminin-binding integrins using monoclonal antibodies against the α or β subunit by flow cytometry. These results indicate that laminin-binding integrins are required for enteric HAdV’s entry and infection of HT-29 cells. Next, we used recombinant PBs produced using the baculovirus expression system to investigate the effect of these soluble PBs on infection of HT-29 cells. Here we observed a decrease in infection of enteric HAdV-41. We used the PB from HAdV-5 as a control but noted that this protein exerted a similar effect on infection as HAdV-41PB. Since, it is known that HAdV-5 uses α3β1 as a co-receptor, which is also a laminin-binding integrin, we speculate the decrease in infection of HAdV-41 with HAdV-5PB could be because of the overlap in integrin usage between the two
HAdV types. We also know that the HAdV-41 GFP vector transduced \( \alpha_5 \) and \( \alpha_{8} \) expressing CHO cells and there is a possibility that the RGD motif containing-HAdV-5PB could inhibit HAdV-41 infection of HT-29 cells by blocking these integrins.

HAdVs use their fiber protein to attach to cells and their PB to enter cells. To establish that the increased transduction of certain CHO-integrin cell lines is not due to the fiber protein, we performed binding assays with the FK protein to CHO-K1, CHO-\( \alpha_2 \), CHO-\( \alpha_3 \) and CHO-\( \alpha_6 \) cell lines. We found that the long fiber knob (LFK) of HAdV-41 bound with equal efficiency to CHO-K1 as the CHO-integrin cells. The same results were obtained for the short fiber knob (SFK) as well as with the whole virus. This data confirms that the fiber proteins do not bind integrins and were not responsible for the increased transduction of CHO-integrin cells with the vector. Thus, we propose that enteric HAdV-41 utilizes laminin-binding integrins for entry and infection and most likely via the PB protein. We also show that integrins function as an accessory receptor or a co-receptor since the LFK blocked binding of HAdV-41 to HT-29 cells, which express CAR – the known receptor for the long fiber. Lastly, we also showed by surface plasmon resonance (SPR) that PB of HAdV-41 interacted with \( \alpha_6\beta_4 \) and the affinity of the interaction was between 30-40 nM while \( \alpha_6\beta_1 \) interacted with a ten-fold lower affinity. Affinities of the interactions between HAdV-5PB and the \( \alpha_6 \) integrins were about three times lower than with the HAdV-41PB.

From the above data, we suggest that the enteric HAdVs have evolved to use a different set of integrins for entry and infection of target cells. Enteric HAdVs have a narrow tropism and are only known to infect the gastrointestinal tract. To the best of our knowledge, expression of \( \alpha_v \) integrins has not been reported in intestinal cells, which makes it plausible that enteric HAdVs make use of other integrins in their infectious cycle. In hindsight, the results and inferences presented in this paper could be justified better had we used human cell lines that over-expressed or lacked the integrins being investigated. HAdVs are
species specific and especially enteric HAdVs grow poorly in cell culture. To be able to transduce CHO cells, we had to use a high MOI of vector, which could have implications on the results. In addition, future experiments will also be validated on cells lines originating from the small intestine such as HIEC-6 and/or HuTu80. A better model system to study the tropism of these viruses could be human intestinal enteroid cultures (276-278) that are derived from stem cells in intestinal crypts. Both undifferentiated and differentiated enteroid cultures can support replication of HAdV-41 (279). Enteroid cultures contain different intestinal cell types such as enterocytes, Paneth cells, goblet cells and enteroendocrine cells. We have shown that enteric HAdVs use laminin-binding integrins as co-receptors. This, however, does not explain why the enteric viruses only infect the gastrointestinal tract since laminin-binding integrins are expressed on many other tissue and cell types as well. Evaluating the integrin profile on the human intestinal enteroid cultures may answer the question about difference in tropism between the different HAdVs. Most likely the determinants of tropism are multi-factorial and depend on interaction and presence of many aspects in the gastrointestinal tract such as enzymes, mucus, bacterial components etc. It will also be beneficial to construct and produce recombinant mutant PBs by substituting potential amino acid residues or by generating whole virions or vectors with mutations in the PB. This may help to pinpoint the exact nature of the interaction with integrins and investigate if mutant PBs can affect entry and infection of host cells. This may also aid in elucidating the integrin-binding site on the PB by performing structural studies such as cryo-EM or X-ray crystallography. Studies of enteric adenovirus tropism and pathogenesis would benefit from a good animal model, but unfortunately there are no such models available.
Paper II
Heparan sulfate functions as cellular receptors for short fibers of enteric human adenoviruses.

In addition to lacking the $\alpha_v$ integrin-binding motif on its PB proteins, enteric HAdVs are also different in that they possess two kinds of fibers – a long and a short. The LF binds CAR but the understanding of SF’s function is incomplete. We have recently shown that the HAdV-41 SFK can act as a toxin (paper IV). We have also shown that HAdV-52 - the only other HAdV with two fibers - uses its SFK as the main ligand and that the SFK binds to Sia-containing glycans on target cells (paper III). The objective with this study was to understand if HAdV-40 and -41 SFs are also engaged in virion binding to cells, and if so, to characterize these interactions with host cell molecules.

In the initial part of the study, we used A549 cells since the HAdV-40 SFK did not bind to the intestinal cells (such as HT-29, HCT-116, COLO205 etc.) that we tested. HAdV-40 SFK and whole virus bind to and infect A549 cells well and we routinely use these cells to cultivate enteric HAdVs. We started out with binding experiments with the SFK of HAdV-40 to characterize the nature of the receptor. Here, we used proteases and a glycolipid biosynthesis inhibitor to examine whether the candidate receptor backbone is a protein or glycolipid. The results indicated that the receptor(s) for HAdV-40SFK contain one or more protein components. To further investigate this interaction (i.e., if it is a protein-protein or protein-glycan interaction), we used enzymes, neuraminidase and heparinase, that degrade known glycan-containing HAdV receptors Sia and HS respectively. 40SFK binding to cells was completely destroyed when the cells were treated with heparinase III, indicating that 40SFK interacts with HS. Next, we pre-incubated the 40SFK with different GAGs and evaluated SFK binding to A549 cells. We used heparin (a soluble analog of HS),
chondroitin sulfate A, a mix of chondroitin sulfate A and C, dermatan sulfate and hyaluronic acid (HA; a non-sulfated GAG). All sulfated GAGs reduced binding of the SFK to cells with heparin being the most powerful inhibitor. HA did not have an affect on SFK binding. These results indicated that the SFK interacted with sulfated GAGs. This was further confirmed when we observed that the SFK’s binding to CHO cells deficient of GAGs was reduced by more than 90% as compared to WT cells. Thereupon, we investigated if heparin, the most potent inhibitor of SFK binding to A549 cells, could also reduce infection of both enteric HAdVs – 40 and 41. Heparin inhibited infection of HAdV-40 and -41 in a dose dependent fashion and at the highest concentration of 100 µM, infection was reduced by 80%. We also observed that a HAdV-41 GFP vector failed to transduce CHO cells lacking HS efficiently. All the above results suggest that the SFK interacts with GAGs and especially with HS, and, that this interaction mediates enteric HAdV binding to and infection of A549 cells. The results also indicate that sulfation is needed for this interaction since the SFK did not interacts with the HA, which has a similar saccharide backbone as HS but lacks sulfate groups.

After establishing that SFK binding is drastically inhibited by heparin at low nanomolar concentrations, we moved on to investigate whole virion binding to A549 cells in the presence of heparin. Unexpectedly, virion binding was not inhibited even at a 1 µM concentration of heparin. These results suggested to us that in the context of a virion, the LFK:CAR interaction takes preference over the SFK:HS interaction, and that attachment of HAdVs to cells is not explained by the presence of the SF and its interactions with HS. Also, unlike HAdV-52, which has equal amounts of LFs and SFs, the LF:SF ratio for HAdV-40 and 41 is 1:6. This implies that the SF should have an important role in the virus’ infectious cycle. Rotavirus, another common enteric virus, has three layers of capsid proteins. The outermost capsid protein (VP4) needs to be proteolytically cleaved to activate the virus particle and allow membrane penetration and infection of cells (280). We hypothesized that, as with rotavirus,
enteric HAdVs might need to be activated for effective infection. Since previous studies have proposed that the SFK makes the virions resistant to low pH (281, 282), a plausible function of the SFK may be to promote binding after low pH activation of the virion. To follow up on this, and on our discovery of a binding partner for the SFK, we examined the binding of synthetic gastric juice-treated HAdV-40 and -41 virions. Here, we used the haploid HAP1 cells, which are easy to modify genetically, to create cells deficient either of CAR or of GAGs. Before adding the treated virions to the cells, we either pre-incubated the cells with CAR-binding (blocking) HAdV-41LFK or treated the virions with soluble heparin. Importantly, binding was reduced by soluble HAdV-41LFK before low pH treatment and by soluble heparin after low pH treatment. The data from these experiments pointed out that HAdV-40 and -41 bind cells via the LF:CAR interaction. However, very interestingly, these viruses bind cells using the SF:HS interaction, after encountering acidic pH. Hence, we put forth a theory in which the first interaction of a virus -coming in through the fecal oral route, thereby being exposed to low pH - with its host cell, is mediated by the SFK:HS interaction. As the virus replicates, large amounts of soluble LFs and SFs are synthesized. These soluble LFs then interact with CAR within tight junctions, which facilitates spread of the newly produced virions and helps in a second cycle of infecting healthy cells through the LFK:CAR interaction.

The enteric HAdVs specifically cause gastrointestinal infections. Earlier studies have contradicting suggestions stating that the SF is responsible for the intestinal tropism (281, 282) but a HAdV-5 vector with HAdV-41 SFs was unable to transduce enterocytes (283). In their study, these chimeric HAdV-5/41 vectors were not subjected to low pH, which may explain their inability to transduce enterocytes. Our results, thus, support that the SF is important for infection of target cells, and in particular during the first round of infection when virions attach cells from the luminal/apical side of polarized enterocytes, where the cell:cell adhesion molecule CAR is usually not exposed.
Studies of respiratory HAdV-5 infection have shown that during infection, fibers are produced in excess, in addition to fibers that are incorporated in virions. These fibers are secreted, and disrupt intercellular CAR:CAR homodimers, which open up the intercellular space and facilitate transmission of subsequently released virions. This may be the case also for enteric adenoviruses. Thus, low pH-activated SF:HS interactions, may explain why HAdV-40 and -41 cause gastrointestinal infection and disease to a much larger extent that other HAdV types.
Paper III
Human adenovirus 52 uses sialic acid-containing glycoproteins and the coxsackie and adenovirus receptor for binding to target cells.

Apart from HAdV types 40 and 41, HAdV-52 is the only other HAdV to possess two fibers. HAdV-52 was isolated relatively recently in 2003 from a small outbreak of gastroenteritis (78). It was classified into a new species – G, since it is dissimilar to the other HAdVs and it is the only human AdV in this species. There have been no studies on the seroprevalence of HAdV-52 in the human population but it has been suggested to be low (78, 284). This feature of HAdV-52 makes it a good vector candidate for gene and cancer therapy. Very little is known about HAdV-52 and how it infects target cells. Thus, the aim of this study was to identify cellular receptors of HAdV-52 to understand the biology of this virus, which would help develop it into a suitable vector candidate.

We first, analysed previously known HAdV receptors and showed that HAdV-52 did not bind to CD46 and did not utilize soluble coagulation factors such as FIX or FX to bind and infect cells. We, however, concluded that HAdV-52 used both CAR and Sia-containing glycans as receptors, since (i) the virus bound to CHO-CAR cells (CHO cells over-expressing CAR) to a significantly greater extent than to control cells lacking human CAR and (ii) the virus could not bind to Lec2 (CHO cells lacking Sia) cells. Treating cells with neuraminidase (enzyme that cleaves Sia) and/or an anti-CAR antibody or treating virus with soluble Sia and/or soluble CAR reduced binding of the virus to cells, which confirmed that HAdV-52 uses CAR and Sia as attachment receptors. Neuraminidase treatment of A549 cells reduced
HAdV-52 infection by 80%. From these results, we could also interpret that the interaction with Sia was more prominent over the interaction with CAR, which indicated that Sia-containing glycans are the chief receptor for HAdV-52. To further characterize the nature of Sia-containing glycans, we treated A549 cells with inhibitors of glycolipid biosynthesis, N-linked glycosylation and O-linked glycosylation. From this, we observed that O-linked glycans are the predominant receptor while glycolipids and N-linked glycans were not of much importance. Next, to investigate which FK was responsible for these interactions, we used recombinant, soluble LFKs and SFKs produced in E.coli. ELISA and SPR results demonstrated that the SFK bound to Sia while the LFK bound to CAR. The affinity of the SFK:Sia interaction was 37 µM and that of the LFK:CAR interaction was in the range of 2.6 to 5 nM, which is similar to the affinities determined for other known FK:CAR interactions. From the crystal structure of the SFK:Sia complex, we discovered the binding sites for Sia on the SFK, which is different than the binding sites for other Sia-utilizing HAdV FKs. There are three Sia binding sites that are formed at the junction of two neighbouring FK monomers making a shallow binding pocket. On mutating single amino acid residues on the SFK that are required for binding Sia, we observed no binding of the mutants on A549 cells or on CHO cells expressing Sia. We validated the correct trimerization of the FK mutants using gas-phase electrophoretic macromolecular analysis (GEMMA), which is a method to determine protein oligomerization. From these results, we predict that the amino acid motif RGN (arginine-glycine-asparagine) is responsible for the interaction with Sia and this motif is not shared by other human HAdVs.

Unlike HAdV-40 and -41, HAdV-52 has equal number of short fibers and long fibers and so the preference in using Sia as the main receptor is surprising since this interaction is weaker than the LFK:CAR interaction. However, CAR is unavailable on the apical side of polarized target cells and the fact that HAdV-37 uses Sia as a functional receptor for infection even though its FK binds with a
stronger affinity to CAR, makes the use of Sia as the prime receptor by HAdV-52 feasible. The interaction between LFK and CAR could function in the disruption of cell-cell adhesion and not only as a cellular attachment receptor. An excess of fiber proteins are made during the production and release of progeny virions, which could bind CAR in the tight junctions and facilitate spread of virions in the tissue (117).
Paper IV

Interaction of human enterochromaffin cells with human enteric adenovirus 41 leads to serotonin release and subsequent activation of enteric glia cells.


Enteric HAdVs infect the gastrointestinal tract and give rise to symptoms such as fever, diarrhea, vomiting and dehydration. Since these viruses are species specific and so far there aren’t any animal models to study disease mechanisms, we set out to investigate the pathogenesis of HAdV-41 using a model system that has been established for studying pathophysiology of rotaviruses. It has previously been shown with rotavirus that on infection of EC cells (an enteroendocrine cell population of the small intestine), serotonin release is triggered causing the activation of vagal afferent nerves. These nerves are connected to parts of the brain that are responsible for nausea, vomiting and diarrhea (270). It was also shown that the rotavirus non-structural protein, NSP4 (271), stimulated serotonin release and has been described as the only known viral enterotoxin today.

EC cells are a population of enteroendocrine cells in the small intestine. They are solely responsible for the production, storage and secretion of serotonin. When EC cells are activated, the granules storing serotonin migrate to the cell membrane and are subsequently secreted. EC cells play an important role in maintaining gut equilibrium with respect to intestinal barrier integrity and motility. Within the intestinal mucosa, EC cells release neurotransmitters (serotonin being one of them) towards the basolateral surface, which is in close proximity to afferent neural endings thus activating them. Underneath the intestinal epithelial cell layer, lies a population of cells
called the EGCs. The EGCs are astrocyte-like and along with neurons form the ENS. Serotonin secretion activates EGCs, which in turn activates the ENS by activating signaling of the extrinsic vagal afferents to the brain. EGCs regulate and maintain intestinal motility and fluid movement across the intestinal epithelium. They also maintain a balance by interacting with both the endocrine and immune systems (265).

The aim of this study was to establish the mechanism by which HAdV-41 induces vomiting and diarrhea, using EC cells and EGCs. Firstly, we determined that HAdV-41 could infect GOT1 cells (a human midgut carcinoid tumor cell line). GOT1 cells express specific EC cell markers (270, 285). We observed that the EC cells infected with HAdV-41 could stimulate serotonin secretion in a dose-dependent manner. We performed a time-kinetic study of HAdV-41 replication in EC cells by qPCR but did not see amplification of viral DNA over the course of 96 hours. However, infecting EC cells for ≤6 hours was sufficient to induce serotonin release, which indicates that infection is adequate to stimulate EC cells and replication is not required. Next, to establish which capsid protein of HAdV-41 was responsible for activation of EC cells; we stimulated the cells with recombinant hexon, PB, LFK and SFK. Here, we noted that only the hexon and SFK were capable of stimulating EC cells to secrete serotonin. These observations were specific for HAdV-41 since HAdV-5 virions (causing respiratory infection in humans) or the FK of HAdV-5 could not stimulate serotonin secretion from EC cells. EC cells express CAR, the receptor for HAdV-41LFK and HAdV-5FK, which proves that the inability of HAdV-5 to stimulate EC cells was not because of the lack of the receptor. To further understand the effect of HAdV-41 stimulation of EC cells, we examined serotonin levels and serotonin-containing granules after infection/stimulation. With confocal microscopy we could ascertain that serotonin levels were depleted in HAdV-41 infected cells due to weaker staining for serotonin but there was increased accumulation of serotonin in granules (or increased granularity), which suggests that infection or
stimulation prepares the EC cells to secrete serotonin from the granules via exocytosis.

To establish if there is communication between HAdV-41-infected EC cells and the EGCs, we tested the activation of EGCs by purified HAdV-41, serotonin or cell supernatant from infected EC cells. EGC activation was determined by the increased intensity of the glia cell activation marker (GFAP) (286-288). We observed that HAdV-41, serotonin and supernatant from HAdV-41-infected EC cells stimulated EGCs. From these results, we propose that stimulation of EC cells and the subsequent activation of EGCs and the ENS due to HAdV-41 infection may be the mechanism by which enteric HAdVs cause vomiting and diarrhea. As described earlier, when HAdVs replicate in host cells, copious amounts of fiber proteins are synthesized, which are not incorporated into progeny virions. In addition to taking part in the infectious cycle of enteric HAdVs, another role of the excess SF proteins could be to stimulate EC cells and cause the physiological effects associated with enteric HAdV gastroenteritis.
Concluding remarks & future perspectives

Gastroenteritis is a very common disease worldwide and viruses are a major etiological agent. HAdVs are the third most frequent cause of viral gastroenteritis and they mainly cause infections in children below five years of age. The HAdV types causing gastroenteritis – 40 and 41, were not well studied with respect to their attachment to and entry of host cells. Another HAdV type that is associated with gastroenteritis is HAdV-52. HAdV-52 is a relatively recent type of HAdV and it was isolated from an outbreak of gastroenteritis. We characterized attachment receptors for this type, which were unknown. Thus, the work in this thesis focused on investigating the functions of two capsid proteins – penton base and fiber.

In the first study, we identified integrins that interact with the PB proteins of HAdV-40 and -41. These PBs are unique, as they do not contain the RGD motif present in all HAdVs, which is needed for virus internalization. We showed that laminin-binding integrins, especially $\alpha_\varepsilon \beta_1$ and $\alpha_\varepsilon \beta_4$, interacted with the PBs of HAdV-40 and -41. For enteric HAdVs, the role of the PB proteins in the viral infectious cycle was not known. Showing that these unique PBs interact with a different group of integrins compared to other HAdVs has added to the knowledge on the function of these distinctive PB proteins.

We also identified cellular receptors for the SFs of HAdV-40 and -41 and for both the fibers of HAdV-52. These are the only three HAdVs that have SFs and their interacting partners and functions were unknown. We showed that the SFs of HAdV-40 and 41 interacted with HS and that the sulfate groups were of importance for this interaction. We also showed that HS was utilized as a receptor only after these viruses encountered low pH conditions as present in the stomach. It was already known that the LFs of HAdV-40 and -41 interacted with CAR and used it for cellular attachment but we
identified CAR as a receptor for the LF of HAdV-52 too. However, most of virion binding was due to the SFs interaction with Sia-containing glycans, which we also identified. We hypothesize that these viruses use their SF receptors for initial attachment. When excess fibers are produced during replication, the LFs bind CAR present in tight junctions and disrupt interaction with CAR present on other cells. Subsequently, this facilitates release of virus from an infected cell and spread to an uninfected cell. For HAdV-40 and -41, the SF may also determine the gastrointestinal tropism of these viruses since the low pH-facilitated interaction between SF and GAGs can only occur when the virions pass through the acidic pH of the stomach.

Finally, we demonstrated an additional function of the SF, where it stimulates enteroendocrine cells (called EC cells) in the small intestine to secrete serotonin. Serotonin secretion in turn leads to the activation of EGCs through which the ENS is activated triggering the vomiting center in the brain. Stimulation of ENS by serotonin secretion and activated EGCs can cause increased intestinal motility leading to diarrhea. Here, we may have discovered a novel viral toxin (after rotavirus NSP4 and HIV-1 Tat proteins) and deciphered a potential disease mechanism of enteric HAdVs. Recently, Hagbom et al. have investigated the function of serotonin receptor inhibitors as treatments to prevent vomiting and diarrhea caused by rotavirus (289). They have shown that treatment with ondansetron (a serotonin receptor inhibitor) can alleviate symptoms of rotavirus infection. It would be very interesting to study the effects of serotonin receptor inhibitors as potential therapeutics for enteric HAdV infections.

Therefore, in my opinion, the results stated in this thesis have helped in the understanding of HAdVs causing gastroenteritis. This increased understanding can be applied in the development of therapy against these viruses or in the development of vectors based on these viruses.

Knowing the functions of the penton base and fiber capsid proteins, in the future, I would like to build-up on this knowledge by determining
whether these proteins are responsible for the narrow gastrointestinal tropism of enteric HAdVs. We now have access to small intestinal cell lines, the suggested target cells for enteric HAdVs, and to human intestinal enteroids. These enteroids are the next best model to animal models since they consist of different intestinal cell populations and mimic \textit{in vivo} intestinal fluid secretion more closely by displaying lumenal swelling. Studying virus-host cell interactions in this system will increase our understanding on the biology of the enteric HAVs manifold.
Acknowledgements

At the end of this five-year journey, I am thinking back to all the new things that I have learnt, encountered and experienced. There are so many people I want to acknowledge and thank for my professional and personal growth.

I will start with my supervisor, **Niklas**, because of whom I came to Umeå. Thank you Niklas for the opportunity to do a PhD in your group. I have learned so much. This experience has strengthened the foundation for my scientific career and I want to strive to do better and make you proud. A big reason for me coming to Umeå was your friendliness and kindness that shone in your emails to me before I accepted the position. You have always been understanding and supportive of me and I have never been afraid or nervous to tell you my feelings about projects or personal life. You are one of the nicest people I have met.

**Lars**, my co-supervisor, I know that I can talk to you about anything under the sun. Thank you for being associated with my projects and for your feedback. Also, thank you for introducing me to archery 😊.

**Kickan**, thank you for all that you have taught me, both personally and professionally. You always had a smile on your face and I always came to you with problems that you mostly always had solutions to. I have missed you this past year in the lab, in the corridors, sharing stories with you, showing pictures of anything and everything and playing plump during Friday fika. Kram!

**Katrine**, after Niklas, you were the first person that I interacted with regarding my PhD position. You have helped me understand how to work and live in Sweden when I was so clueless in the beginning. Your presence and all that you took care of for us at Virology has helped a great deal and I am so very grateful to you. **Lena, Maja** and **Veronica** thank you for your help with administrative work and for always finding time to answer my questions.

My colleagues in Niklas’ group - **Naresh**, thank you, thank you and thank you. You are usually the first person I run to in the group when
I need help or when I just need to talk to someone. You have helped me so much throughout. We started our PhD journeys and lives in Sweden at the same time and have dealt with similar problems and experienced similar joys. Thank you for everything. I wish you all the best for the future and hope you achieve everything your heart desires. **Annasara** and **Stefan**, I am so glad that I agreed to share office with you. From science to life, I love sharing my thoughts and concerns and experiences with both of you and I have grown to like you more as friends than colleagues over the years. I will miss the leg pulling, the bet making, the never-ending who-reaches-the-office-first arguments and all the other silly and serious talk that goes on in the pink office. I will miss you both a lot when I leave. **David** and **Nitte**, you have always been approachable and so helpful for anything that I have needed. David, you have swooped in and saved my projects or my issues with Illustrator or my computer woes so many times. Thank you!

I would like to acknowledge everyone who has been in the group in the past. **Raquel, Carin, Rickard S, Victor, Sharvani, Marija, Lijo, Marko, Mari** and **Agnieszka**, I feel I had a personal connection with each one of you. You added to the enjoyable working and social atmosphere of the group. Raquel, I am so glad we are still in touch. I wish you and Antonio all the best with all your travels and with life. Lots of love to you. Carin, thank you for being the sweet, caring and understanding person that you are. You have taught me many aspects regarding the lab during my initial PhD days and you have always been patient and kind. Rickard, thank you for guiding me in the beginning and giving me valuable tips to survive this PhD journey. Victor, you were an amazingly fun office-mate! Marija, I will treasure our trip to Abisko and hopefully see you in Leiden 😊.

**Elin**, it was a pleasure and joy to supervise you for three summers. I learnt to be a supervisor with you and I hope you liked your time and lab experience at Virology and thank you for all your hard work and help in my projects.

**Sarah**, you came to the lab as a student but left as a very, very dear friend. I absolutely enjoyed and loved and still cherish the time you spent here and all the things we did together. It was short but it felt like I had known you forever 😍. I miss you!
To all my colleagues (past and present) at Virology, thank you for adding to the pleasant work environment and helping me around the department. Chaitanya, Arun and Koushikul, we have shared so much together. I could go on and on about everything but for now I just say thank you for being there for me! Richard, you are amazing, kind and so sweet. I am lucky to have you as my friend and it means a lot to me when you just stop by my office or lab or in the corridor and ask me how I am doing. Kirstin, I am so glad that I got to know you. Thank you for the bursts of positive energy during my thesis writing! Emma and Katarina, I will miss our silly lunch conversations. Watch out for me in the next winter Olympics! I would like to acknowledge everyone that I have played plump with at Friday fika. It was sheer joy and I miss this tradition now!

The friends that I made outside of Virology thank you for enriching my life. Mridula, Shi Pey, Swetlana and Geetanjali you all are special because you were my first girlfriends in Umeå. We have done so much together. Thank you for the all the good times and the wonderful company. All the best with everything in the future! Gautam, Tanushree and Christophe, I am so happy to have gotten to know you. We had a short overlap in our times in Umeå but we made the best out of it. Steffi and Mahesh, Radha and Damini thank you for all the times we have spent together and all the memories we have made together (there have been loads of it!) I am very grateful for your friendship. Chinmay, thank you for all the lovely meals you have cooked for me, all the (memory) games that we played together and all the 4 am cricket matches that you let me watch at your place! Anubhav, Laxmi, Sanat and Rupali I have enjoyed the times that we have gotten together.

All my pre-Umeå friends, who have always supported, encouraged and loved me. I hate being so far away from you guys but I am so glad that we still know what is happening in each other lives. Ancy, Anisha, Rashmi, Shamik, Nikita, Sindhu, Ketaki, Madhura, Shashi, Finny and Prashant I love you and I love you and I love you!

I would also like to thank my grandmother who has always wished for me to become a ‘big scientist’. I know that you are proud of me and I am so grateful for your blessings. Mamma, hope you feel better soon and become your usual cheerful self again.
Fredrik, my love, thank you for all that you have taught me both in
the lab and in life. I admire your approach to and outlook on life. I
aspire to learn some of that from you. Thank you for bearing with me
when I got unbearable. Thank you for listening when I vented or
ranted. Thank you for the solutions to my problems. As I have always
said, you were my best and worst student. I have learnt to tackle
science and my projects differently (and in a better way) from you.
Thank you for everything! Also thank you for the cover and the only
non-English section of this thesis 😊. I would also like to thank
Fredrik’s friends with whom we have had fun get-togethers and board
game evenings. Also, thank you to Fredrik’s family in Helsingborg
who have hosted me and welcomed me warmly into their homes and
family.

Lastly, I would like to thank my family. Thank you Amma and Papa.
I am what I am because of your unconditionally support and
encouragement. Bad lab days ended not-so-bad because I would see
your smiling faces on Skype. You both are a phone call away.
Anything I need, anything I do, any help, any problem, sharing small
joys or big sorrows, you are there for me and with me. How many
ever times I say thank you or show my appreciation will never be
enough. Chechi and PJ, my other pillars of strength. I cannot imagine
life without you. Chechi, you know what you mean to me. I lack
words to express how grateful I am for everything that you have done
for me and taught me. I look up to you and I am so proud of you. PJ,
after my thesis defence, I will take a long break and read the books
and watch the movies you have recommended 😊. Watching Neel
grow up has given me so much joy and happiness. Neelkutti, you are
the cutest, smartest, loveliest little boy I know and I feel privileged to
be your Cherri. I love you all so, so, so, very, very, very, much and I
miss you more than anything!
Bibliography


25. T. C. Mast 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).


239. A. E. Tollefsen et al., The adenovirus death protein (E3-11.6K) is required at very late stages of infection for efficient cell lysis and release of adenovirus from infected cells. *J Virol* 70, 2296-2306 (1996).


262. J. J. Short et al., Substitution of adenovirus serotype 3 hexon onto a serotype 5 oncolytic adenovirus reduces factor X binding, decreases liver tropism, and improves antitumor efficacy. Mol Cancer Ther 9, 2536-2544 (2010).


266. V. Erspamer, B. Asero, Identification of enteramine, the specific hormone of the enterochromaffin cell system, as 5-hydroxytryptamine. Nature 169, 800-801 (1952).


274. H. Ido et al., Laminin isoforms containing the gamma3 chain are unable to bind to integrins due to the absence of the glutamic acid residue conserved in the C-terminal regions of the gamma1 and gamma2 chains. J Biol Chem 283, 28149-28157 (2008).


