Cryo-electron microscopy of HPV16 pseudovirions reveal changes in capsid conformation upon furin cleavage

---

Melissa Lauren Marx

MRXMEL007

Supervisor:

Dr. Georgia Schäfer

Co-supervisor:

Dr. Jeremy Woodward

Presented for the Degree

Master of Science in Medicine

Medical Biochemistry

Division of Medical Biochemistry and Structural Biology

Department of Integrated Biomedical Sciences

Faculty of Health Sciences

University of Cape Town

21 December 2020

Declaration

I, Melissa Lauren Marx, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Date: 21 December 2020

Acknowledgements

This has been a challenging but rewarding two years and would not have been possible without my amazing support systems, both academic and personal. All of you have helped me get through several trials during my MSc, most notably a very extended lockdown and several rounds of nation-wide power-cuts!

I am incredibly grateful to my two stellar supervisors, Dr Georgia Schäfer and Dr Jeremy Woodward. Jeremy, without your technical expertise and outstanding problem solving skills, many aspects of this project would not be possible. I am so grateful for all your help and patience, and to have worked with you for the past two years. Georgia, thank you so much for the opportunity to join your lab with such an interesting project, and for all your help, advice, and patience along the way, I very much appreciate it! I would like to thank both my supervisors for the opportunity to attend several conferences, which have helped my development as a postgraduate student tremendously.

I would also like to thank all the staff within the Electron Microscopy Unit, especially Mohammed Jaffer, who was always available when I needed help with the electron microscopes. Thank you to Prof. Sewell, Michael Woodward, Andy Rabagliati, Miranda Waldron, Sean Karriem and Nasheeta Hanief for all your assistance these past few years, it does not go unnoticed! To Graham Christians, Xolani Nonzinyana and Roshan Ebrahim, thank you for all your hard work in the lab and office, and for keeping the lab running so smoothly. To the staff at eBIC (Diamond Light Source, UK) including Dr James Gilbert and Dr Alistair Siebert, thank you for all your patience and guidance with the cryo-EM data collection sessions.

I am also grateful for all the moral support from the students and post-doctoral fellows which I have had the opportunity to work with. Dr Mulelu, Dr Lubbe, Dr Sarron and Dr Blumenthal, thank you for all the advice and reassurance. To Sinead Carse, Lenye Dlamini, and all the other postgraduate students I have met along the way, thank you for making my postgraduate studies so enjoyable.

To my family and friends, thank you for the unwavering support throughout the years. To my parents, Caroline and Llewellyn Marx, thank you for everything you have given me, especially your love and support along the way, I am forever grateful. To my brother and sister, thank you for the moral support!

I would also like to give my sincerest thanks to the Poliomyelitis Research Foundation, Dr Georgia Schäfer, the National Research Foundation and the GCRF Synchrotron Techniques for African Research Technologies (GCRF-START) for funding my studies. Without your financial support, this research project would not be possible, and I would not be here today, and so for that I am forever grateful. Lastly, I would like to thank GCRF START, Dr Jeremy Woodward and Dr Georgia Schäfer for the opportunity and funding to conduct a research visit to eBIC (Diamond Light Source; UK), Collecting data and touring at a world-class facility like eBIC has been a life-changing experience, and I will be eternally grateful to all of you who helped make that possible.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

3D	Three-dimensional
°C	Degrees Celsius
AIDS	Acquired immunodeficiency syndrome
BCA	Bicinchoninic acid
BSA	Bovine serum albumin
BSL	Biosafety level
BPV	Bovine papillomavirus
CCD	Charge coupled device
CTF	Contrast Transfer Function
Cs	Spherical aberration
CsCl	Caesium chloride
C-terminus	Carboxyl terminus
dH_2O	Distilled water
DMEM	Dulbecco's modified eagle's medium
DNA	Deoxyribonucleic acid
DQE	Detector Quantum Efficiency
eBIC	Electron Bio-Imaging Centre
ECM	Extracellular matrix
E. coli	*Escherichia coli*
EGFR	Epidermal growth factor receptor
EM	Electron microscopy
EMDB	Electron Microscopy Data Bank

FSC	Fourier shell correlation
GAG	Glycosaminoglycans
G1	Gap 1 phase
GLuc	Gaussia luciferase
GUI	Graphical user interface
GPU	Graphical processing unit
HA tag	Haemagglutinin tag
HEK293TT	Human embryonic kidney 293TT cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
HPV	Human papillomavirus
HPV16-PsV	Human papillomavirus type 16 pseudovirion
HSB	High-salt buffer
HSPGs	Heparan Sulphate proteoglycans
Å	Angstrom
$e^-/Å^2$	Electrons per angstrom^2
kb	Kilobase
KCl	Potassium chloride
kDa	Kilo Daltons
KLK8	Kallikrein-8
kV	kilo Volts
L1	Late HPV protein 1
L1/L2	Late HPV protein 1 and 2

L2	Late HPV protein 2
L	Litre
-m	-meter
m-	Milli
M	Molar
$MgCl_2$	Magnesium chloride
MPI	Message passing interface
MW	Molecular Weight
N	Nano
NaCl	Sodium chloride
ND10	Nuclear domain 10
nm	nanometers
N-terminus	Amino terminus
ORF	Open-reading frame
ori	Origin of replication
PDB	Protein data bank
PBS	Phosphate buffered saline
PML	Promyelocytic leukaemia
PsV	Pseudovirion
PVs	Papillomaviruses
RAM	Random access memory
Rb	Retinoblastoma
RCF	Relative centrifugal force

RELION	Regularised Likelihood OptimisatioN
RMSD	Root mean squared deviation
RPM	Revolutions per minute
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SEM	Standard error of the mean
SNR	Signal to Noise Ratio
SPIDER	System for Processing Image Data from Electron microscopy and Related fields
SV40	Simian Virus 40
TEMED	Tetramethylethylenediamine
TEM	Transmission Electron Microscope/Microscopy
TRIS	Tris(hydroxymethyl)aminomethane
S phase	Synthesis phase
U	Units
UCSF	University of California, San Francisco
VLPs	Virus-like particles
Vol	Volume
v/v	Volume/volume
w/w	Weight/weight
α	Alpha
β	Beta
γ	Gamma
µ	Micro-
~	Approximately

Abstract

Persistent infection by oncogenic human papillomavirus (HPV) is the primary cause of cervical cancer, a leading cause of cancer deaths in women worldwide. There are no treatments for HPV infection, and although prophylactic vaccines are effective and safe, they are HPV type specific, provide little therapeutic benefit and developing countries often have limited access to these. Therefore, additional measures against HPV infection are urgently needed.

Preventing HPV entry into host cells is an attractive option for therapeutic intervention. The HPV capsid is icosahedral and consists of two proteins, L1 and L2, which participate in entry and infection of host cells. During entry, the virus capsid attaches to the cell surface via binding to heparan sulphate proteoglycans (HSPGs). Cleavage of L2 by a host protease, furin, is necessary for infection and is thought to facilitate a conformational change in the virus capsid. Furin cleavage may affect the ability of HPV to bind to sulphated glycoproteins and a HSPG substitute, heparin.

Understanding these proposed structural changes may aid in the development of therapeutics targeting virus entry. Here, we directly visualize the conformation changes to HPV16 pseudovirions (HPV16 PsVs) resulting from cleavage of L2 by exogenous furin using cryo-electron microscopy (cryo-EM).

At 5 Å resolution, we observed that furin-cleaved HPV16 PsVs capsids display widespread changes in the arrangement of capsomeres relative to uncleaved control virions. This structural change is relevant because heparin has previously been observed to bind to the HPV16 capsid in the canyon surrounding the capsomere at the five-fold icosahedral symmetry axis, but not in other canyons between capsomeres, related by pseudo-symmetry. This suggests that differences in the relative orientations of the surrounding capsomeres to each other either prevent or allow heparin binding. We observed a narrowing of the putative heparin binding site by 0.4 Å after furin cleavage and propose that this change may be responsible for the transfer of HPV from cell-surface HSPGs to the unknown entry receptor(s) by a yet unidentified mechanism.

1 Introduction

1.1 Human papillomavirus and cancer

Papillomaviruses (PVs) are a group of viruses infecting diverse species such as reptiles, birds and humans (1, 2). The human papillomaviruses (HPV) are non-enveloped, double-stranded DNA viruses, infecting the mucosal and cutaneous epithelia (2-6). HPV infections a are highly prevalent sexually transmitted infection and can be classified as low- or high-risk infections, based on their oncogenic potential (7, 8). Infection with HPV types of the beta genus are considered low-risk as these subtypes cause benign warts and asymptomatic lesions in the skin and mucosal epithelia. Although considered low-risk, HPV types 5 and 8 within this genus are correlated with the development of cutaneous squamous cell carcinoma, while persistent infections with any or a combination of the twelve high-risk HPV types within the alpha genus have been linked to several cancers (5-8). High-risk types HPV16 and HPV18 are present in 72% of HPV associated cancers, making these two HPV types the most common types associated with oncogenesis (5, 9-11). As HPV16 is present in many HPV associated cancers and is widely studied, HPV16 is the focus of this investigation.

HPV linked cancers include oropharyngeal, head, and neck cancer and most notably, cervical cancer (5, 12-18). Almost all cervical cancer cases are attributed to high-risk HPV infection and cervical cancer is the fourth most diagnosed cancer in women worldwide (15, 19, 20). Cervical cancer is a leading cause of cancer deaths in women in 42 countries, the majority of which are in sub-Saharan Africa, including South Africa (5, 20). The high incidence and mortality of cervical cancer in sub-Saharan Africa may be linked to the human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) epidemic within the region coupled with under-resourced vaccination and cervical cancer screening programs (9, 21-23). Indeed, HPV infection is a necessary but insufficient cause of cervical cancer as persistent infection over several years as well as co-factors (such as HIV/AIDS induced immunosuppression) is needed for possible cancer development (5, 7, 8, 17, 20).

Sexually active individuals are likely to become infected with one or more HPV type(s) during their lifetime, but most of these infections are asymptomatic and can be cleared by the immune system within one to two years (18, 24). However, immune clearance is slower for HPV16 and related high-risk types, increasing the risk that infection with these viruses will lead to cancer (section 1.4) (11, 24). Other co-factors for cervical cancer include high HPV viral load, long-term oral contraceptive use, tobacco smoking, multiple births, and chronic inflammation. Patient genetic and immunological factors are likewise significant in determining the outcome of HPV infection. Co-infection with HIV, Herpes simplex virus 2, or *Chlamydia trachomatis*

infection may also increase the risk of cervical cancer (2, 10, 22, 25-27). As persistent high-risk HPV infection is a key component of cervical cancer development, preventing HPV infection may thwart cervical cancer development. Therefore, cervical cancer is potentially an entirely avoidable condition. Highly effective and well-tolerated HPV vaccines are available but have some shortcomings, despite substantial efforts by vaccination programs (section 1.2). Consequently, cervical cancer remains a leading cause of cancer deaths in women worldwide (20). This makes the development of additional methods to prevent HPV infection of critical importance for cervical cancer prevention.

1.2 Current HPV preventative strategies

Presently, there are no treatments for HPV infection, but prophylactic vaccines trademarked as Cervarix™, Gardasil® and Gardasil-9® exist to prevent infection in unexposed individuals. (28-35). All three vaccines prevent infection by the most common oncogenic types HPV16 and HPV18. Gardasil® and Gardasil-9® provide additional protection against the non-oncogenic HPV types, HPV6 and HPV11 (9, 31, 32, 35). Gardasil-9® provides added protection against the oncogenic HPV types 31, 33, 45, 52 and 58 (34, 36). HPV vaccines function by allowing the immune system to create antibodies against the HPV type-specific L1 capsid protein, eliminating HPV after exposure (33, 37). This process is effective and precise, generating antibodies specific to the HPV types vaccinated against (33). For example, in the UK, HPV16 and 18 prevalence decreased more than 6% in 16-to-18 year old females and more than 10% in 19-to-21 year old females after vaccination with Cervarix™, only six years after widespread vaccination schemes were implemented (38). In the United States of America, there was a decrease in the incidence of cervical precancerous lesions of more than 21% in women who had received at least one dose of an HPV vaccine, only six years after vaccine introduction (39). These results illustrate that prophylactic vaccination is an effective strategy to prevent HPV infection and associated cervical pre-cancers. Importantly, cancer prevention strategies are more cost-effective and less invasive than harsh cervical cancer treatment options (40). However, these vaccines are not therapeutic and are ineffective against existing cervical lesions in HPV infected women (32). Moreover, long-term HPV vaccine efficacy is unknown, making screening and other cancer prevention strategies still necessary (10, 40).

There are other notable shortcomings of the currently available HPV vaccines, most notably limited HPV type cross-protection. As the vaccine target protein (the major capsid protein L1) differs across HPV types, immune responses to HPV infection are only effective against the specific types vaccinated against, rather than a broad cross-protection across different HPV types (32, 33, 36, 41). This is especially problematic for individuals at risk for infection by an

oncogenic HPV type not currently covered by available vaccines and immuno-compromised individuals such as HIV positive individuals, as they are at increased risk of HPV persistence and associated cancers (21). Non-vaccine HPV types like HPV35, HPV45 and HPV54 are prevalent in Africa, which also has a high HIV/AIDS burden. In a study of women in Nigeria, Ghana and South Africa, 97.1% of HIV positive women were also infected with at least one HPV type, or several HPV types concurrently including non-vaccine types. These studies also emphasise the importance of cervical cancer prevention in HIV positive individuals (21, 42, 43). However, vaccine uptake is surprisingly low, even in countries with reasonable healthcare infrastructure. For example, only 39.7% of all teen girls in the United States of America have completed a vaccination regimen (10). In South Africa, school-based vaccinations have been implemented by the National Department of Health since 2014 and have been reasonably successful, but coverage across resource-limited populations is still poor (44, 45).

Limited vaccine uptake may be partly explained by cost and accessibility, as the HPV vaccines are relatively expensive and require repeat administration for maximum protection. Additionally, HPV vaccines are not thermostable, requiring cold storage before use, which is difficult to achieve in low resource settings present in rural areas of the world (10, 32, 33, 36). This creates a difficult situation for many individuals, who may not be able to afford HPV vaccination or are without access to well-resourced vaccination facilities. Another factor affecting vaccine uptake, especially in South Africa is reliable knowledge on HPV, as rural South African communities are largely unaware of HPV as a risk factor for cervical cancer. Cultural stigmas against vaccinating for sexually transmitted diseases may also affect vaccine uptake (36).

Although the available HPV vaccines are safe and effective for type-specific HPV infections, their various shortcomings highlight the need for development of alternative therapies targeting HPV infection. Possible therapeutic targets include the entry mechanisms used by the virus during infection of host cells. However, changes to the HPV capsid during entry, HPV receptor use and details of cell attachment entry are insufficiently understood. Therefore, it is the focus of the present investigation, to enhance knowledge of structural changes to the HPV capsid during entry, as detailed knowledge of proposed changes in capsid conformation during entry may aid in the development of small-molecule inhibitors of HPV infection.

1.3 HPV capsid structure and genome

The HPV capsid is a robust protein shell which encloses and protects viral DNA (Figure 1.1[A]) (46). The capsid is ~55 nm in diameter, non-enveloped and icosahedral, made up of 360

copies of the major capsid protein L1 (6, 47, 48). Each L1 protein forms a chain (chains A to F; Figure 1.1 [B, C]. Five L1 protein chains make up a single capsomere, known as a 'pentamer', (Figure 1.1 [B]), and 72 pentamers make up the HPV capsid (4, 48, 49). These pentamers can be arranged in two ways on the HPV capsid. Pentavalent pentamers are made up of five L1 chain A's and arranged with one pentamer surrounded by five other pentamers, on the icosahedral five-fold axes of symmetry. Hexavalent pentamers are made up of five L1 chains B to F, and one hexavalent pentamer is surrounded by six other pentamers (Figure 1.1 [B]). The spaces between capsomeres form 'canyons' within the HPV capsid (Figure 1.1 [A]) (4, 48, 49).

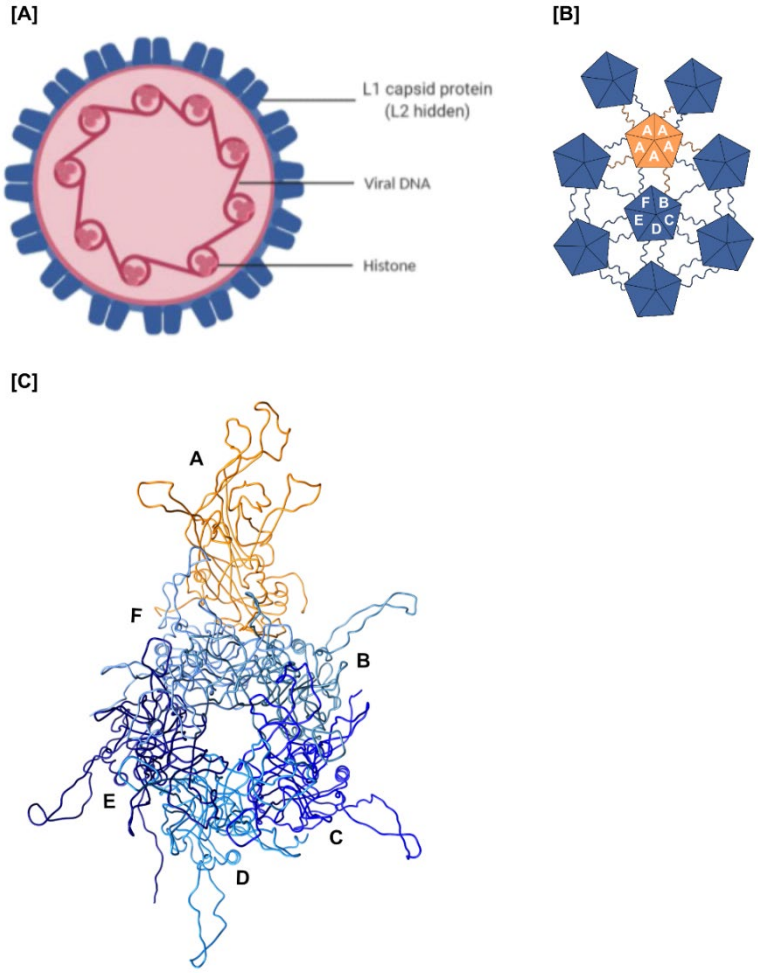


Figure 1.1. Graphic representation of the HPV16 capsid, genome, capsomere arrangement and asymmetric unit. **[A]** Schematic of the papillomavirus capsid and viral genome with associated histones as labelled. **[B]** Arrangements of capsomeres on the HPV capsid. Capsomeres are indicated by pentagons made up of five L1 proteins. Hexavalent capsomeres are indicated in blue, while a pentavalent capsomere is shown in orange. Each of the five L1 chain within a capsomere is indicated by a triangle and labelled with letters. The pentavalent capsomere is made up of five L1 chains labelled chain A, while each hexavalent capsomere is made up of five L1 chains labelled chain B, C, D, E, F. The flexible carboxyl terminus arms are indicated by coloured lines linking the capsomeres. The orange carboxyl terminal arms originate from the pentavalent capsomere, while the blue carboxyl terminal arms originate from the hexavalent capsomere. **[C]** The L1 only asymmetric unit (PDB ID: 5kep) of the HPV16 capsid is shown, with each (of six) L1 chains labelled as chains A to F. Chain A forms part of the pentavalent capsomere and is indicated in orange, while chains B to F, form a hexavalent capsomere are coloured in different shades of blue (1, 50).

L1 is ~55 kDa in size and can self-assemble into the full viral capsid, even in the absence of the second (minor) HPV capsid protein, L2 (2, 47). The five L1 chains within each pentamer form a central β jelly roll (Figure 1.2). The carboxyl terminus (C-terminus) of each L1 protein (Figure 1.1 [B], Figure 1.2 [B] and [C]) extends outwards to neighbouring capsomeres, through spaces between adjacent capsomeres (4, 50). Each L1 protein also contains flexible surface-exposed loops (Figure 1.2 [B] and [C]), which are often the site of antibody binding and are poorly conserved across papillomavirus species (51, 52).

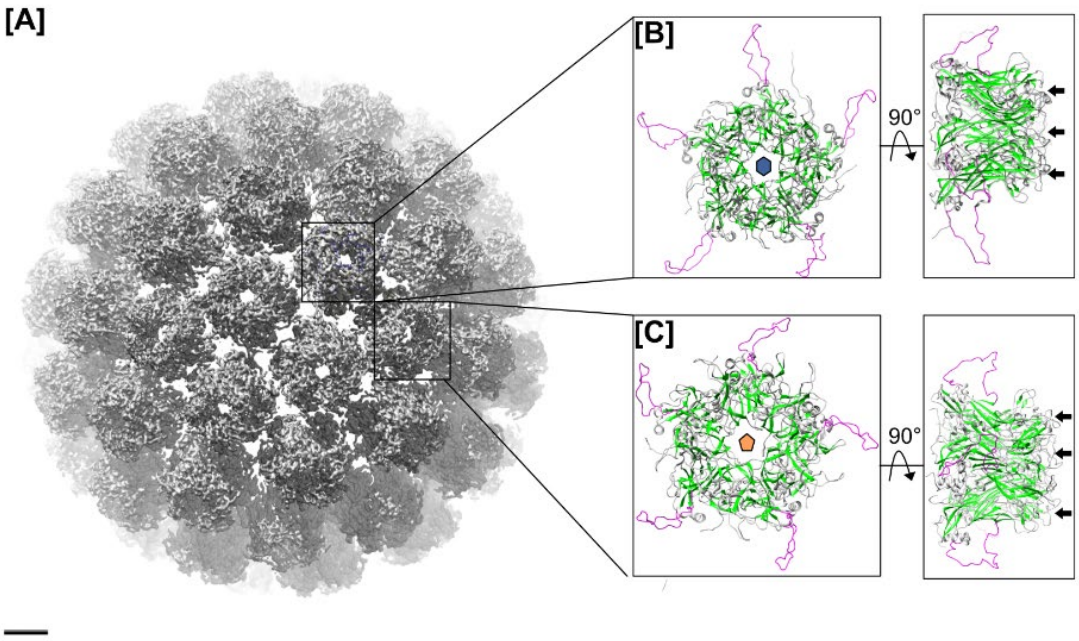


Figure 1.2 The structure of HPV16. [A] Cryo-electron microscopy density map of HPV16. Scale bar 10 nm. **[B]** structure of the HPV16 L1 hexavalent pentamer (PDB ID: 5kep) indicated by a blue hexagon. Each L1 carboxyl terminus is indicated in magenta, the central β barrel is shown in green. **[C]** Structure of the HPV16 L1 pentavalent pentamer (PDB ID: 5kep) indicated by an orange pentagon. Each L1 carboxyl terminus is indicated in magenta, the central β barrel is shown in green, while the location of surface exposed loops of each L1 chain are indicated by arrows (50).

Aside from forming the HPV capsid, the L1 protein facilitates initial attachment and entry into host keratinocytes (section 1.5). The minor capsid protein L2, is hidden from the HPV capsid surface, although part of the amino terminus (N-terminus) is exposed on the capsid surface (52). The position and the exact number of L2 proteins per HPV capsid is inconclusive. Evidence suggests that L2 occurs at $1/30^{th}$ of the abundance of L1, however, some researchers suggest 36-72 copies of L2 per HPV capsid (36, 53, 54). Antibody binding studies suggest that each L1 pentamer interacts with a single molecule of L2 (i.e., 72 copies of L2 per virion), although other biochemical studies suggest a stoichiometric ratio of 5:1, with a total of 12 L2 proteins associated with 60 of the L1 molecules (54-56). The number of L2 molecules per HPV capsid may differ between HPV expression systems, as the L2 gene is often transcribed from a codon optimised plasmid with varying expression levels.

Cryo-electron microscopy (Cryo-EM) analysis suggests that L2 is located at the base of each capsomere, (Figure 1.3 [A]). L2 apparently fills the central cavity of each L1 pentamer and about sixty N-terminus amino acids may be exposed on the capsid surface (52, 57-59). Two other hypotheses exist for the location of L2: one is that L2 is spread throughout the capsid within both hexavalent and pentavalent capsomeres, (Figure 1.3 [B]) and the second is that L2 is located only at a unique capsomere within the capsid (50, 55, 60). L2 is important for viral infection within the host cell (sections 1.4 and 1.6) (61-63). L2 also promotes late gene expression and viral DNA encapsulation during virion assembly, leading to infectious HPV capsids (53, 64). However, L2 does not appear to form disulphide bonds within the HPV capsid and the interactions between L1 and L2 are likely non-covalent and hydrophobic (52, 54, 65, 66). This contrasts with several disulphide bonds present within/between the L1 proteins. To form the HPV16 capsid, disulphide bonds, especially those formed by cysteines 175, 185 and 428 in the L1 protein, stabilise individual capsomeres, while disulphide bonds formed by other L1 cysteines like cysteine 229 and 379 assist in maintaining overall capsid structure and increasing resistance to proteolysis (4, 50, 51, 56, 65, 67, 68). Site-directed mutagenesis also suggests that lysine 54, arginine 74, and lysine 82 are needed for successful HPV16 PsV assembly (69). Intra-capsomere bonds, forming the pentamers, are more highly conserved than intercapsomere bonds, and may be more essential for the capsids' function (48).

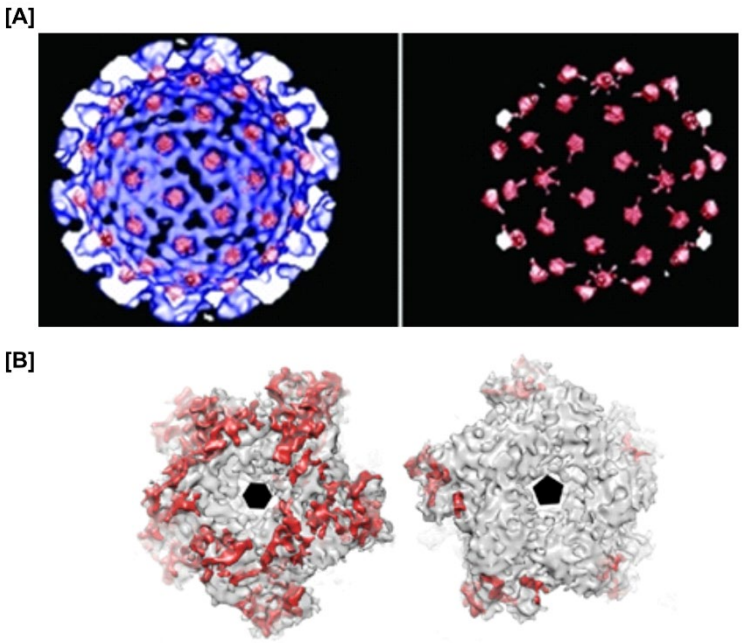


Figure 1.3 The possible locations of the HPV16 L2 minor capsid protein. The HPV L1-only cryo-electron microscopy map was subtracted from the L1 and L2 containing HPV cryo-electron microscopy maps, and putative locations of L2 are indicated as red density. **[A]** Cryo-electron microscopy map indicates L2 is located on the interior of each capsomere. Left panel shows L2 density (red) within the internal view of HPV16 capsid (blue). Right panel indicates the L2 density alone (55). **[B]** High resolution of HPV16 capsomeres (grey) indicate L2 density (red) is present primarily around the hexavalent capsomere, with some L2 around the pentavalent capsomer. Hexavalent capsomeres are indicated by the hexagon, while pentavalent capsomers are indicated with the pentagon (50).

Intercapsomere interactions in papillomaviruses are also stabilised by an ‘invading arm’ model, where the L1 C-terminus extends into an adjacent capsomere, forms a loop, and returns to the ‘donor’ capsomere (Figure 1.1[B]). The invading portion of the loop runs along the floor and up the walls of the canyons between capsomeres. The disulphide bonds formed by cysteines 175 and 428 occur at the ‘elbow’ of the invading loop. The L1 N-terminus also participates in stabilising the capsomeres, mostly by interactions with other L1 N-terminus extensions (4, 50, 51, 56)**.** However, as the HPV capsid is made up of both pentavalent and hexavalent pentamers (Figure 1.1 [B] and Figure 1.2 [B, C]), not all intercapsomere contacts are identical (48, 50). The intercapsomere contacts between hexavalent and pentavalent capsomeres are symmetrical, where a single L1 chain on the pentavalent capsomere (Figure 1.1 [B]) only contacts a single L1 chain within a neighbouring hexavalent capsomere. However, the contacts between two hexavalent capsomeres are dissimilar. Three hexavalent capsomere L1 chains contact four neighbouring capsomeres, while the other two L1 chains within the hexavalent capsomere contact the remaining two surrounding capsomeres (Figure 1.1 [B]) (48, 50).

The HPV capsid encloses an ~8 kilobase (kb) double stranded DNA genome, which is associated with host histone proteins (Figure 1.1 [A] and Figure 1.4) (2, 51). The viral genome encodes six ‘early’ (E) regulatory proteins and two structural ‘late’ (L) proteins, within eight open reading frames (ORFs) (Figure 1.4) (2, 36, 56). These early and late open reading frames are separated by an upstream regulatory region (URR) containing the viral origin of replication, which is responsible for the regulation of gene expression, genome replication and packaging into progeny virus capsids (2, 70, 71). Polyadenylation signals (A_L and A_E) for viral genome replication are also indicated (Figure 1.4).

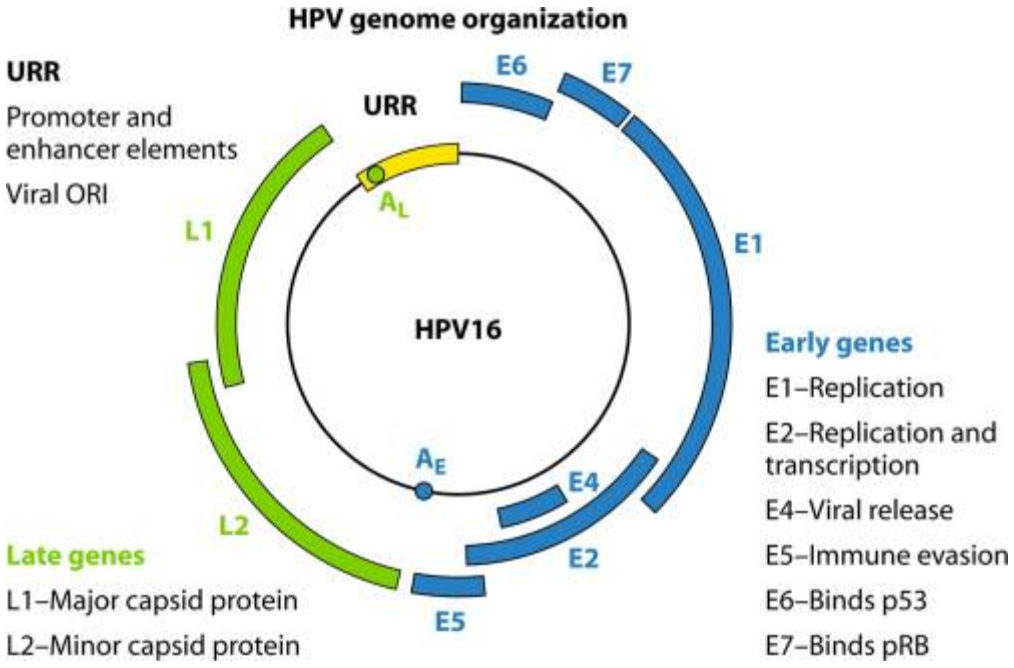


Figure 1.4 Diagram illustrating the organization of the HPV16 genome and a summary of each open reading frames' (ORFs') function, as described in the text. The genome contains early (E) and late (L) regions. The upstream regulatory region (URR) regulates early ORF transcription. A_L and A_E are polyadenylation signals used during virus transcription (3).

The early HPV proteins include E1 and E2, which are involved in viral genome transcription and replication, while the E5, E6 and E7 proteins are considered oncogenic and facilitate malignant host cell transformation (section 1.4) (12, 18, 70, 72). E1 has DNA origin of replication (ori) binding and unwinding capabilities, while E2 is a transcription activator. These two viral proteins form a complex for enhanced viral DNA replication (2, 72). E2 also has a role in viral genome expression, by regulating transcription of the HPV E6 and E7 proteins. During viral DNA integration into the host genome, this regulation is lost, leading to overexpression of the E6 and E7 viral oncoproteins (18). E6 and E7 are directly associated with oncogenesis, as they inactivate two essential tumour suppressor proteins, p53 and retinoblastoma protein (pRB) respectively (section 1.4) (2, 3, 18). E4 has a role in viral genome amplification, expression of capsid proteins, and virus release (73). The role of E5 in virus replication is unclear, although E5 and E6 are involved in immune system evasion (section 1.4).

The late proteins of HPV are L1 and L2, which form the icosahedral viral capsid and are important during the early steps of host-virus interactions (section 1.5) (71, 74). L2 is also important for viral infection within the host cell, facilitating endosomal escape and transport of the viral genome to the cell nucleus (sections 1.4 and 1.6) (61-63). Each viral ORF and resulting protein successfully functions to allow viral replication and progeny virus assembly. This process is highly effective, as the virus alters the host cell machinery to allow greater

virion replication. However, this has unintended consequences, enabling metaplastic transformation of the host cells, which, if undetected by the host immune system, may cause malignant tumorigenesis within the HPV infected tissue.

1.4 HPV infection cycle and malignant transformation of host cells

The duration of the HPV infection cycle is anywhere from a few weeks to months and is dependent on the cell cycling and differentiation of host keratinocytes, although the regulation mechanism is unclear (2, 3, 70, 75). In HPV uninfected cervical epithelial tissue, basal keratinocytes divide along the basement membrane, while keratinocyte maturation occurs vertically, without undergoing further cell division (Figure 1.5, left). HPV is presumed to enter the epithelium through micro-abrasions or wounds within the tissue, which expose the basal keratinocytes. Infection of basal keratinocytes likely happens while the cells are migrating and dividing to repair damage to the epithelium. After successful infection of basal keratinocytes, an initial round of viral DNA replication occurs, which maintains the HPV genome as a low copy number episome at around fifty to one hundred copies per keratinocyte. This represents a latent HPV infection, with no phenotypic changes to the host cell (2, 3, 76, 77). After HPV infection, expression of viral proteins causes infected cells to divide vertically, delaying cell maturation and epithelial differentiation, while increasing both viral genome replication and HPV protein expression (Figure 1.5, right) (2, 71). During progression to cervical cancer, the HPV genome is integrated into the host cell genome. This is assumed to be non-random, enabling increased cell cycling with the potential for malignant cell transformation by destabilising the host genome (2, 70, 77). Moreover, loss of the E2 ORF, during viral genome integration, causes dysregulation of the viral proteins, E6 and E7, resulting in chromosomal instability and gene mutations within the host cell genome, further increasing the likelihood of oncogenesis (76-78). Interestingly, high-risk HPV types are associated with more genome deletions, which may contribute to their greater oncogenic potential (78).

HPV uses several evasion tactics to avoid detection and elimination by the host immune system. For instance, the HPV lifecycle occurs almost entirely within the infected host cell, thereby restricting expression of HPV proteins within the underlying epithelial tissue. Furthermore, HPV proteins do not cause cell lysis, making antigen detection by the immune system difficult. The coupling of the HPV lifecycle to the short lifecycle of host keratinocytes further limits the time for HPV detection by the immune system. Moreover, mature virions are released from the uppermost layers of the infected epithelium, far away from the immune cells circulating within the lower epithelial layers (17, 24, 56, 74, 79-82). Viral proteins may also directly affect aspects of pathogen detection by the immune system. E6 downregulates

expression of human leukocyte antigen 1 (HLA-1) and E5 also decreases monocyte recruitment to the site of infection, allowing the virus to avoid elimination in HPV infected cells (2, 83). The E6 protein may also play a role in host immune evasion, by downregulating signalling molecules required for interleukin-6 and interleukin-8 expression and type 1 interferon expression within infected host cells. This prevents recruitment of natural killer cells to the HPV infected cells and cell death (17, 18, 82). These immune system evasion strategies employed by HPV allow the virus to persist in tissues for long periods (24).

Once the HPV infected cell enters the upper epithelial layers and exits the cell cycle, there is an increase in copies of the HPV genome per cell, with increased expression of viral genes and proteins (Figure 1.5; right) (2, 3, 77). Viral DNA replication is dependent on the host cell, as the virus only encodes a single DNA replication enzyme, E1. Consequently, other HPV proteins function to induce viral DNA synthesis, inhibit cell growth regulation signals, prevent apoptosis, and delay host cell differentiation (2, 3, 76). Although a host cell's progression to a malignant state is a non-productive infection for HPV, the overexpression of HPV proteins and the proteins' effects on infected keratinocytes aids in malignant transformation of cells (70, 78). Central to oncogenesis are the viral proteins E6 and E7. E6 binds the host protein p53, a key cell protein for the regulation of the cell cycle, DNA repair and apoptosis, leading to cell cycle disruption and promotion of cell proliferation. E6 also prevents apoptosis, by interfering with the activities of pro-apoptotic proteins and preventing telomere shortening, extending the lifespan of the infected cell. This may increase the likelihood of spontaneous gene mutations, which accumulate and lead to tumorigenesis (2, 3, 18). The E7 oncoprotein binds to unphosphorylated cellular retinoblastoma (Rb) protein, which overrides the G1/S checkpoint of the cell cycle, leading to increased cell division (3). E7 also interacts with cell proteins like p21, which allows the infected cell to bypass cell cycle checkpoints, leading to cell proliferation and potentially oncogenesis (2, 3, 18, 84).

Although the expression of viral proteins causes increased cell cycling with the potential for carcinogenesis in infected tissue, the primary function of the viral proteins is to enable the increased production of progeny virions within infected host cells. During progeny virion assembly, the HPV genome replicates in the nucleus and L1 and L2 capsid proteins are assembled in the cell cytoplasm. L2 then localises into promyelocytic leukaemia (PML) bodies within infected cells before recruitment of L1 proteins and viral DNA with cellular histones, assembling into the native HPV, This progeny virus is then released from the most superficial epithelial layers (Figure 1.5, right) (53, 64, 70, 85). After release from the epithelium, progeny virions must undergo the complex mechanisms to enter epithelial cells for successful infection.

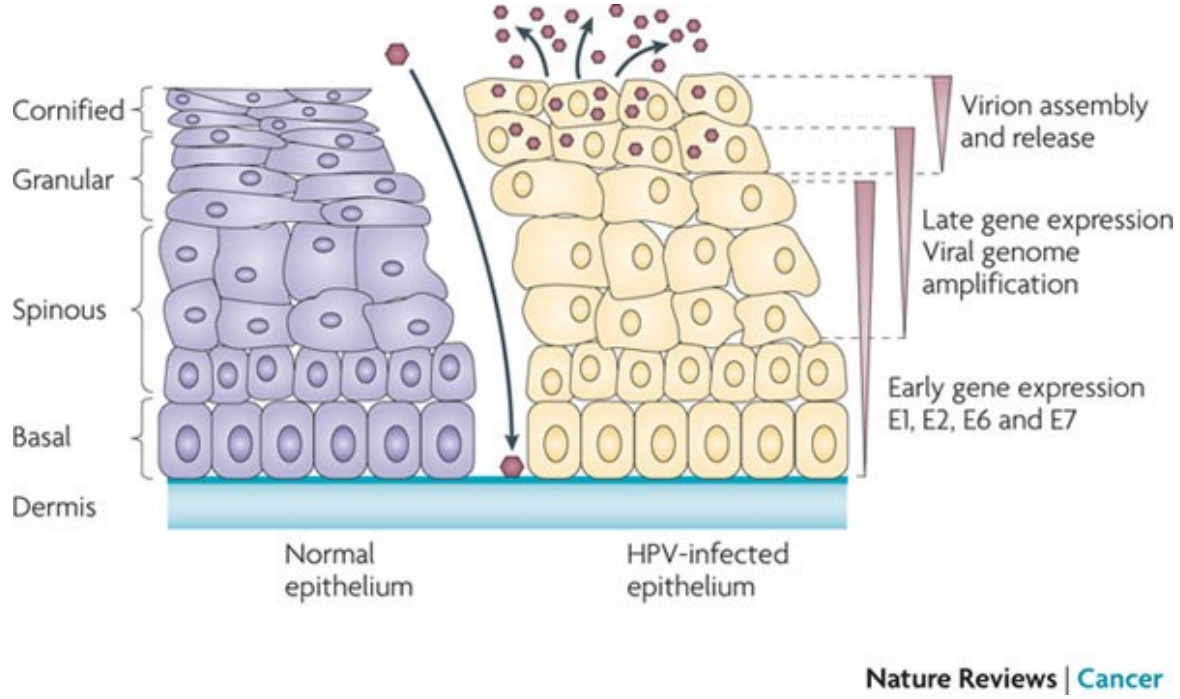


Figure 1.5. Diagram of the HPV lifecycle in the epithelium. The layers of the epithelium are indicated on the far left. HPV (red hexagon) infects basal keratinocytes exposed through mico-abrasions, indicated by the gap between the epithelial layers and the arrow. HPV uninfected epithelium is shown on the left ('normal' epithelium; purple) and HPV infected epithelium is on the right ('HPV infected epithelium'; yellow). The lifecycle of HPV is indicated in brief on the right, beginning with early gene expression within the basal epithelium, followed by late gene expression and viral genome amplification in the upper layers. Finally, the progeny virion is assembled and shed in the upper epithelial layers (86).

1.5 HPV transmission and entry mechanisms

HPV is spread by direct skin-to-skin contact and enters the basal epithelium through micro-abrasions in the tissue (Figure 1.5), initially binding to the basement membrane followed by transfer to the basal keratinocyte surface (17, 80, 87). The different HPV types are epitheliotropic, only successfully infecting mitotically active basal keratinocytes in the epithelium of the skin, the genital and oral mucosa (2, 3, 6, 88). This strict cell tropism is suspected to be dependent on events following HPV entry into cells, as HPV5, a non-genital HPV type can infect genital tissues but less effectively than either HPV16 or 18, which infect genital tissues (6, 74, 89). The limited cell tropism is also closely linked to keratinocyte differentiation (section 1.4 and Figure 1.5) (2, 36, 80, 85).

Like many viruses, HPV may use a co-ordinated multi-receptor entry process, initially binding to a low-affinity receptor present on many cell types, before transfer to a more specific entry receptor (47, 90). Papillomaviruses likely use an evolutionarily conserved binding receptor or

receptor complex, as bovine papillomaviruses (BPV) and HPV compete for binding to the cell surface, while Simian Virus 40 (SV40), does not interfere with HPV binding (59, 89). Early studies reported a specific HPV attachment receptor with many copies on the cell surface, and clusters of virus-like particles (VLPs) were observed attached to HeLa cells, rather than VLPs binding evenly to all areas of the cell surface. VLP binding increased until apparent receptor saturation, as increasing the amount of VLPs available did not affect further VLP binding (59, 89). Importantly, HPV binding was found to be trypsin sensitive, indicating a membrane protein may be used by HPV for cell attachment (89).

1.5.1 Heparan sulphate proteoglycans as the initial cell surface receptor for HPV

Heparan sulphate proteoglycans (HSPGs) on the basement membrane and keratinocyte surface are the likely initial cell binding receptor for most oncogenic HPV types *in vitro* and *in vivo*, including HPV16 and 18 (47, 53, 90-92). HSPGs are complex, negatively charged glycoproteins containing covalently bound heparan sulphate chains. These heterogenous molecules are present on the cell surface and extracellular matrix (ECM), and act as receptors for cytokines, chemokines, and growth factors (93, 94). They are also receptors for selective cellular proteases, regulating the activity of these enzymes and their distribution in the ECM. Therefore, HSPGs are involved in numerous cell processes, including wound healing, cell signalling, cell motility and endocytosis (6, 93, 94). HSPGs are expressed in high numbers on epithelial cells, the natural host cell for HPV (6, 69, 90, 95, 96). Moreover, many other viruses make use of HPSGs as an initial non-specific cell binding receptor, and HPV seems to be no exception to this (94, 96-98). Since protein binding to HSPGs on the cell surface prevents protein degradation by extracellular proteases, HPV attachment to HSPGs may also function to inhibit virus degradation prior to entry. HPV may also use HSPGs to mediate proteolytic processing by specific proteases necessary for successful entry (section 1.5.3) (94, 99).

To demonstrate the involvement of HSPGs in HPV binding, inhibition of cell glycosaminoglycan (GAG) synthesis or virus incubation with cells deficient in cell surface GAGs was found to reduce HPV VLP binding to cells (47, 92, 96, 100). Treatment of HaCaT (spontaneously transformed human keratinocytes), COS-7 (green monkey derived fibroblast-like cells) and primary human keratinocyte cells with heparinase and sodium chlorate (to disrupt cell surface heparan sulphate structure and charge) drastically reduced HPV11, 16, 31 and 33 VLP binding (47, 79, 95, 96, 100, 101). Similarly, HPV11 VLP binding to HaCaT and COS-7 cells was inhibited when the VLPs were incubated with HSPG binding competitors, including heparin (section 1.5.1.1) and protamine, suggesting the virions bound to these

HSPG binding competitors, rather than the cell surface HSPGs. A compound binding to heparan sulphates, a N,N′-bisheteryl derivative of dispirotripiperazine (DSTP27), prevented HPV16, 18 and BPV1 infection of transformed human embryonic kidney (HEK) 293TT cells (47, 79, 92). Moreover, x-ray crystallography and cryo-electron microscopy studies have directly indicated that HPV16 and HPV18 bind to sulphated glycoproteins (50, 102, 103). These studies all indicate a role for cell surface HSPGs for HPV binding during infection. However, HPSGs as the initial HPV binding receptor has also been challenged, as HPV31 virions derived from organotypic raft culture bound to cells independently of HSPG expression (104). Likewise, HPV16, 18, 31 and 45 organotypic raft culture derived virions had differing dependencies on cell surface sulphated GAGs. Alternatively, the virions could make use of several HSPGs on the basement membrane and/or cell surface to mediate infection (41, 102, 103). In line with this hypothesis, Johnson *et al.* found that HPV16 must bind to HSPGs present on both the basement membrane and cell surface within the female murine genital tract for successful infection *in vivo*, indicating HPV binds to several different HSPGs on the cell surface and basement membrane during a natural infection (6).

The exact HSPG type used by HPV during entry is unknown and may differ depending on the infection system used (6). Current estimates are that somewhere between $1X10^4$ and $2X10^4$ HPV entry receptor molecules are present on epithelial cells, while there are many more ($\sim10^6$) proteoglycans of different types on the cell surface (47, 89, 96, 105). Candidate HSPGs for HPV entry are the cell membrane bound HSPGs (syndecans and glypicans) or the secreted ECM HSPGs (agrins and perlecans) (94). Syndecan-1 and syndecan-4 are 31-45 kDa glycoproteins with covalently attached chains of heparan sulphate and/or chondroitin sulphate. Glypican-1 and glypican-4 are two 57-69 kDa glycoproteins with covalently attached heparan sulphate chains. These HSPGs are all expressed on the extracellular matrix of epithelial cells and fibroblasts, cell types known to facilitate productive HPV infections (47, 94).

To investigate which HSPG facilitates HPV16 VLP binding, keratinocytes were incubated with increasing heparinase concentration to remove cell surface HSPGs, or phosphatidylinositol-specific phospholipase C (PI-PLC) to specifically remove glypicans. Heparinase reduced VLP binding in a dose dependent manner, while PI-PLC did not, implying glypicans are not the HSPG type used during VLP binding to keratinocytes (69, 100). Syndecan-1 may be the primary HSPG used by HPV16 VLPs for infection *in vitro*, as cells deficient in most cell surface HSPGs, but expressing transfected syndecan-1 had the greatest HPV VLP binding capacity, compared to cells transfected with either syndecan-4 or glypican-1. Moreover, syndecan-1 was the primary HPV binding HSPGs on keratinocytes infected with HPV11 and BPV (47, 100). Syndecan-1 is also associated with infection by other pathogens and is a co-receptor for

many of the growth factor receptors that HPV makes use of during infection (section 1.5.3), making this HSPG type the likely HPV receptor (94, 106).

However, other research suggests that the different HPV types have a more complex dependency on cell surface glycosaminoglycans than described above. In the presence of chondroitin sulphate, HPV18 virions were non-infectious, while HPV16, 31 and 45 were able to infect cells. Furthermore, adding exogenous carrageenan (a sulphated polysaccharide derived from algae) did not inhibit HaCaT cell infection with HPV16 and HPV31, but did inhibit infection with HPV18 or HPV45 (107). Potentially, HPV binds to more than a single HSPG type and recruits other HSPGs for binding during a sequential entry process, or there may be redundancy in the HSPG used by HPV during entry (6, 41, 90, 96, 102, 103, 107). Another possible explanation for the interaction between HPV and cell surface HSPGs is that HPV may not dissociate from HSPGs on the cell surface during entry but instead the capsid becomes coated with HSPG shed from the cell surface, followed by subsequent internalisation of the HSPG coated capsid (108).

1.5.1.1 Heparin binds to HPV experimentally but is not involved in HPV infection

As HSPGs are long and heterogenous molecules which are difficult to synthesise, heparin (a related anionic sulphated polysaccharide) is often used to study the interaction between HSPGs and HPV during infection. Notably, heparin is not involved as a cell binding receptor in the natural HPV infection but is used as an *in vitro* substitute for HSPGs (6, 41, 50, 79, 94, 102, 103). 93% of HPV11 VLPs can be eluted by free heparin, and HPV types 16, 33, and 39 were found to bind efficiently to a heparin-bovine serum albumin (BSA) complex compared to control BSA. Adding exogenous heparin had an inhibitory effect on HPV11 and 16s' ability to infect HaCaT and COS-7 cells (47, 79, 96, 101). A similar inhibition was observed for HPV 16, 18, 31, 45, 58, 59 and 68 pseudovirion infection of COS-7 cells (95). Low concentrations of heparin (0.05 mg/mL) abolished infectivity of HPV16 and HPV33 pseudovirions, whereas this inhibitory effect was not seen with other glycosaminoglycans (96). When heparin was added to HaCaT cells before adding pseudoviruses, HPV binding did not decrease (101). This implies that heparin obstructs HPV binding to HSPGs by binding to the virus capsid, rather than the cell surface. Increasing concentrations of heparin decreased HPV VLP binding to cells in a dose-dependent manner, with 50% inhibition observed at the lowest heparin concentration (0.25 g/ml) and complete inhibition at 32 g/ml heparin, similar to results obtained with HPV11 VLPs (47, 69, 100). Together, these studies indicate that several HPV types bind to heparin, and heparin competes with the cell surface HSPGs for binding sites on the HPV capsid. This

confirms the suitability of heparin as a substitute for HSPGs when studying HPV entry mechanisms.

However, HPV's sensitivity to heparin may differ depending on the HPV production system used. When organotypic raft-culture derived HPV was used to bind to or infect HaCaT cells, the various HPV types had different responses to the presence of heparin. HPV16 was able to attach to the cell surface in the presence of heparin, while HPV31, an evolutionarily related HPV type was unable to bind to the cell surface. Likewise, HPV18 and a related type, HPV45, were unable to bind HaCaT cells (107). These results suggest that HPV16 may make use of a non-HSPG receptor, while other HPV types make use of sulphated polysaccharides for binding. However, the viruses had dissimilar responses to infection in the presence of heparin. HPV16 and HPV31 infection of HaCaT cells was unaffected by heparin. Similarly, HPV45 infection was resistant to heparin, while a related type, HPV18, was inhibited by heparin (107). The discrepancy in the effect of heparin between organotypic derived or pseudovirion production HPV particles may be explained by the presence of cellular factors and proteases, which are necessary for infection (see 1.5.2), being present in the differentiating tissue culture, while being mostly absent from the pseudovirus system. Accordingly, the differentiating tissue derived virions may be processed by these cellular factors prior to cell infection, allowing them to bypass the initial HSPG attachment step during infection and resulting in resistance to heparin binding (87, 99, 104, 106, 109, 110). Interestingly, mutations in L1 surface exposed amino acids on the HPV16 capsid implicated in cell surface binding (section 1.5.1.2) display decreased cell binding when incubated with heparin, suggesting that specific amino acids allow HPV binding to HSPGs on the cell surface (69).

1.5.1.2 Surface exposed amino acids of L1 mediate HPV attachment to heparan sulphate proteoglycans

HPV attachment to HSPGs (or heparin) may be facilitated by the L1 viral capsid protein, as the C-terminus of L1 in HPV16 and HPV11 interacts with glycosaminoglycans, and neutralising antibodies against L1 inhibited HPV16 VLP's attachment to keratinocytes (47, 50, 69, 79, 100, 105, 111). Two of these antibodies, H16.V5 (V5) and H16.U4 (U4), bind to different areas of L1 HPV16, thereby preventing infection. V5 binds to a conformational epitope at the top of the L1 pentamers and prevents virus attachment to the underlying extracellular matrix (ECM) but HPV may still bind cell surface HSPGs. Similarly, an x-ray crystallography structure of the L1 only pentamer, complexed with heparin found that heparin only bound to surface-exposed regions of the pentamer, despite all regions being physically accessible for binding (102). U4 recognizes a linear epitope within the C-terminal arm of L1

that forms the “invading arm” and permits virus binding to the ECM while blocking binding to cell surface HSPGs (105, 111, 112). Consequently, the L1 regions bound by these antibodies identified either on the surface of the L1 proteins, or within the L1 C-terminus “invading arm” are necessary for HPV binding to HSPGs on the cell surface.

The L1 C-terminus is well conserved among HPV types and is characterized by a cluster of six to eight basic amino acids within the final fifteen residues and a shorter three to four basic amino acid cluster upstream. The putative heparin or HSPG binding motif is BBBXB where B is a basic amino acid (lysine or arginine) residue, and X is any amino acid. Potentially the L1 C-terminus in the canyon between capsomeres is the site for HSPG binding (47, 52, 69, 94) (94). Site-directed mutagenesis indicated that lysines 278, 356, and 361 on L1 were essential for both cell surface binding and infectivity with HPV16 PsVs composed of both L1 and L2 capsid proteins (69). A combination of these lysine mutations had a greater effect than single mutations, while the triple mutant had the least binding and infection capacity, suggesting that these lysines bind HPV to the cell surface in a co-operative manner (69).

However, mutation of arginine amino acids in the vicinity of these three lysines had little effect on HPV16 PsV binding and infectivity, implying HPV cell binding and infectivity was dependent on charge distribution within the binding area, rather than a specific amino acid sequence (47, 69). Structural studies indicated that lysines 54, 56 and 356 with threonine 266 and arginine 285 from two neighbouring L1 monomers bind heparin. Additionally, heparin was also found to bind asparagines 57 and 450 and lysines 59, 443 and 452 in another binding patch on the pentamer surface. This study also demonstrated that heparin (and potentially HSPGs) binds to either positively charged or polar residues, which would allow charge-charge interactions and hydrogen bonds to form with the anionic heparin molecule(s) (90, 102, 103, 113, 114). However, during HPV33 infection, the L1 C-terminus was not required, as a peptide composed of HPV33 L1 C-terminus amino acids did not inhibit HPV33 infection, even at a high concentration (5 mg/mL). Furthermore, antiserum directed against this peptide had no effect and HPV33 particles without some L1 C-terminus amino acids still bound to heparin-coated ELISA plates, suggesting that the L1 protein of HPV33 is not involved in cell surface binding (96). Moreover, heparan sulphate molecules inhibited HPV binding and infection, but other proteoglycan molecules with similar numbers of negatively charged residues did not inhibit HPV infection, suggesting the interaction between HSPGs and virions could depend on structural features, rather than electrostatic interactions on HSPGs (93). Further studies will be needed to investigate whether this applies to other HPV types too.

The minor L2 capsid protein does not seem to be involved in initial cell surface binding as BPV VLPs, composed of either the L1 capsid protein or both capsid proteins (L1 and L2) were

equally effective at inhibiting native BPV binding. Additionally, L1-only HPV33 VLPs still bound to HeLa cells, without the need for the L2 protein (59, 89). Antisera against L2 did not affect the ability of BPV to bind to cells but affected infectivity and L1 only PsVs were much less infectious than L1/L2 HPV pseudovirions. This research implies that L2 is not involved in cell surface binding but is involved in later stages of infection (58, 59).

1.5.2 Conformational changes to the HPV capsid during entry

As detailed above, HPVs are suggested to use HSPGs as an initial non-specific contact to the host cell, which is followed by virus conformational changes and transfer to an entry receptor or receptor complex. Several proteins are proposed as necessary for productive infection, some of which may facilitate changes in the HPV capsid structure, including cyclophilin B, kallikrein-8 and, most notably, furin (section 1.5.3). Furin is a 96 kDa multidomain proprotein convertase, with a vital responsibility for converting proteins including integrins, neuropeptides, growth factors and hormones into their mature forms, which then participate in homeostasis, cell signalling, and embryogenesis (115-118). Furin contains a highly negatively charged catalytic site, which cleaves protein precursors with the amino acid sequence – arginine/lysine-X-X- arginine/lysine where the first and last amino acid can either be arginine or lysine, and X is any amino acid. The HPV L2 capsid protein contains a furin cleavage site on the N-terminus, at arginine 12, which is highly conserved across papillomavirus species (106, 110, 116, 118).

This furin cleavage site was demonstrated to be functionally relevant in HPV16 and BPV1 infection, as mutated arginine amino acid residues resulted in reduced HPV infection levels (106). Furin is widely expressed in mammalian tissues and is present both intracellularly, in the trans-Golgi network, on the cell surface and extracellularly (87, 115, 116, 118, 119). These properties allow the enzyme to be exploited by pathogens during entry to enhance infectivity and pathogenicity, including bacteria like *Bacillus anthracis*, *Clostridium septicum* and many viruses, such as avian influenza, HIV-1, semliki forest virus, equine herpesviruses, cytomegaloviruses, ebolaviruses and SARS-CoV2. The glycoproteins of these viruses must be processed by furin to form mature virions and to enhance infectivity (97, 115, 118-125). While furin is generally membrane-bound, an active form can be secreted, but it is presently unknown whether furin is membrane-bound or the soluble when the enzyme acts on HPV16 particles during natural HPV infection (105). However soluble furin can cleave the L2 minor capsid protein, making this form of furin sufficient for infection *in vitro* (105, 126).

Proteolytic processing of L2 by furin is a rate-limiting step during HPV entry, indicating the importance of this step for HPV infection (105). If L2 is not cleaved during entry, L2 and the viral genome fail to escape the endosome during intracellular trafficking (section 1.6) and the

virus is non-infectious (106). Treatment of HPV5, 6, 16, 18, 31 and BPV PsVs with an inhibitor of furin and related proteases prevented infection in HeLa cells. This was not the case with inhibitors of other proteases, indicating furin or a furin-like protease was necessary for infection (106). To confirm this finding, HPV16 PsVs and BPV1 virions were incubated with two furin deficient cell lines: FD11 (Chinese Hamster Ovary (CHO)) cells, and LoVo (colon carcinoma) cells. These cells were entirely resistant to either papillomavirus infection, even though the cells may express other non-furin proprotein convertases (106). Similarly, HPV16 PsVs which were incubated with exogenous furin prior to cell infection do not need cellular furin for infection *in vitro* or *in vivo*, hence exogenous furin successfully processed the virus capsid for entry (99, 105, 126, 127). This was further confirmed by incubating either furin processed or furin unprocessed virions in the presence of a furin inhibitor before cell infection. The furin inhibitor only prevented infection of the furin unprocessed virions, presumably by preventing the proteolytic activity of cellular furin on these HPV capsids (105, 126, 127). Moreover, HPV16 PsVs are much more infectious when cleaved by exogenous furin, compared to uncleaved HPV, demonstrating the vital role of furin in enhancing the infectivity of HPV (126).

Furin likely acts on the HPV capsid at the cell surface as inhibition of several endosomal and lysosomal proteases, had little effect on infection, while *in vivo,* cleavage of L2 may occur on the basement membrane (87, 106, 112). Furin processing is thought to occur near the exterior of L1 pentamers on the HPV capsid, as an L1 specific antibody (H16.V5) which binds to surface loops on the L1 pentamer, blocked furin cleavage of L2 (112). In contrast, another L1 specific antibody (H16.U4), which binds to 'invading arms' around HPV capsomeres (magenta in Figure 1.2) had little effect on L2 cleavage, suggesting that the furin cleavage site, comprising part of the L2 N-terminus, is unlikely to be located in the 'canyons' between capsomeres (112). Interestingly, an L2 specific antibody enhanced furin cleavage of L2. Potentially, the antibody prevents its' L2 epitope from being buried after antibody binding, increasing furin's' accessibility to the cleavage site on L2 (112). Furin appears to incompletely cleave HPV L2 as several research groups have demonstrated about 50% cleaved L2, with some uncleaved L2 detected within the late endosome of infected cells (106, 112, 126, 127). However, it is unknown at present whether this percentage of virions is completely furin cleaved within the total HPV population, or whether all the HPV L2 molecules are partially cleaved across the entire virus population (112). This situation may represent a natural infection, with heterogenous HPV capsids having differing degrees of exposure of the L2 N-terminus and susceptibility to furin cleavage (112, 127). Even though L2 is partially cleaved by furin, there is evidence that the HPV capsid undergoes changes to the capsid quaternary structure to enable successful infection (87, 112, 127, 128).

This conformational change after furin cleavage has been demonstrated indirectly by antibody binding studies using the RG-1 antibody, which detects the amino acids 17-36 on the L2 protein, which are adjacent to the furin cleavage site. L2 exposure of the N-terminus was assessed by RG-1 detection of a human influenza hemagglutinin (HA) tag inserted at the extreme N-terminus of the L2 protein. Antibody binding to this HA-tagged N-terminus of the L2 HPV protein increased over time, indicating greater exposure of the L2 N-terminus while RG-1 is unable to bind to furin uncleaved HPV16. Although the L2 N-terminus is partially exposed on the capsid surface (52), this result suggests these amino acids only become surface exposed after furin cleavage. Immunofluorescence staining also demonstrated that the L2 amino acids 17-36 were exposed over time after HPV16 PsV binding to the basement membrane (105, 127, 128). This suspected conformational change was confirmed by immunoprecipitation, where the RG-1 antibody was able to bind furin cleaved HPV16 particles, but not furin uncleaved particles (87, 127). Likewise, HPV capsids which have undergone a maturation step are in a different capsid conformation compared to immature capsids and are resistant to furin cleavage. This implies that the furin cleavage site is unavailable until the viral capsid is in a conformation similar to that of immature virions (106). Furthermore, treatment of HPV16 PsVs bound to HaCaT cells with a furin inhibitor prevented exposure of the RG-1 antibody binding site on L2 (128).

Furin mediated exposure of the L2 protein is necessary for internalisation and intracellular infection steps after cell surface binding (58, 106, 110, 129, 130). Full length HPV16 L2 capsid proteins weakly bound to HeLa cells, while a peptide made up HPV16 L2 protein residues 13 to 31 had greater binding to cervical cancer cell lines (131). None of the other peptides made up of other regions of the L2 amino acid sequence bound HeLa cells, emphasising the importance of the region near the furin cleavage site for cell surface binding (131). However, other studies have found that HPV16 L2 peptides made up of the first twelve amino acids in L2, adjacent to the furin cleavage site, did not inhibit HPV16 PsV infectivity of COS-1 cells (58). COS-1 cells are non-human and non-keratinocyte cells and may not represent the most appropriate cell line for assessing HPV16 infectivity, given the strict virus cell tropism. Importantly, it cannot be excluded that these L2 peptides are internalised by cell endocytosis, rather than the internalisation pathways used by HPV. Interestingly, heparinase treatment of HeLa cells reduced L1-only HPV16 VLP binding, but not the L2 peptide binding, suggesting that HPV16 uses the L1 protein to attach to HSPGs on the cell surface, before L2 binding to a non-HSPG receptor (131). However, HPV16 L1/L2 virions were unable to bind heparinase treated HeLa cells, suggesting that the region of the L2 peptide which bound to HeLa cells is inaccessible for initial binding (131).

Currently, it is proposed that changes to the HPV capsid after interaction with host molecules needed for infection, occurs mainly to the L1 capsid protein, given that this is the predominant capsid protein and conceals the bulk of the L2 minor capsid protein (50, 55, 60). It is likely that L2 also undergoes some structural rearrangement after these processing steps, but this is difficult to identify as the location and the exact number of L2 molecules per HPV capsid is inconclusive. Indeed, two L1 specific antibodies which bind conformational epitopes on the L1 protein (H16.V5 and H16.E70) prevented exposure of the RG-1 L2 antibody binding site on HPV16 PsV capsids, suggesting the L1 capsid protein occludes the RG-1 antibody binding site. This is unlikely to be due to steric hindrance between the two antibodies, as cell bound HPV16 virions were able to bind both L1 and L2 specific antibodies. Consequently, the two L1 specific antibodies prevent the cell surface conformational change necessary for furin cleavage and L2 antibody binding site (128).

The conformation change to L1 may also be caused by a keratinocyte secreted enzyme, serine protease kallikrein 8 (KLK8) which processes L1 before cell entry (132). Similarly, another host protein, cyclophilin B may cause a conformational change in the viral capsid, although the effects of cyclophilin B on HPV entry is debated. Potentially cyclophilin B is essential for intracellular infection steps rather than HPV entry (53, 105, 112, 133). HPV33 L1-only and L1/L2 VLPs were incubated with two antibodies specifically recognising L1 epitopes on capsomeres, H33.B6 and H33.J3. The antibody-coated VLPs were then exposed to heparin. H33.B6 impaired VLP binding to heparin while H33.J3 did not impair VLP-heparin contact, suggesting these antibodies prevent HPV33 infection through different mechanisms (93). H33.J3 neutralised cell bound but not free virions, suggesting the H33.J3 epitope is only available on HPV33 after cell attachment and a capsid conformation change. Interestingly, this antibody has similar characteristics to a BPV antibody, which binds to hexavalent capsomeres, neutralising infection without affecting cell binding (93). Potentially after cell surface binding, a conformation change within the hexavalent capsomere allows H33.J3 binding to the HPV capsid. This is somewhat likely, as HPV and BPV capsid proteins are functionally conserved (131).

The proteolytic steps by host molecules and conformational changes to the HPV capsid are thought to be necessary for HPV transfer to the unknown secondary receptor or receptor complex (section 1.5.3), and viral entry into host cells (6, 87, 99, 101, 110). The process of HPV entry into host cells is very slow and this may be due to the number of conformational changes and interactions with the host cell and associated proteins for successful entry (59, 105). However, the effects of furin cleavage on the HPV capsid and the proposed capsid conformation changes cannot be seen by negative staining and low-resolution electron

microscopy (EM) (126, 127). Consequently, high-resolution imaging of furin cleaved particles is necessary to directly determine whether furin cleavage of L2 leads to conformation changes in the HPV capsid, as proposed in the literature.

1.5.3 HPV co-receptors for binding and entry

HPV pseudovirions' interaction with the cell surface changes from heparin sensitive to heparin resistant, which may indicate the presence of a secondary receptor (96). Other evidence for the presence of a second, more specific entry receptor is that combined inhibition of two molecules proposed as HPV binding and/or entry receptors, heparan sulphate and laminin 332, did not completely abolish HPV16 PsV infection of HaCaT cells (92). Additionally, HPV16 PsVs were unable to infect pgsA-745 cells, which are deficient in cell surface HSPGs. However, these cells could be infected when exogenous furin, a protease necessary for infection (sections 1.5.2 and 1.5.3) was added to the virions, allowing these virions to bypass the HSPGs and enter the cell through unknown mechanisms. As the virions could infect cells in the absence of HSPGs, the presence of a second, non-HSPG receptor responsible for HPV uptake was proposed (99). Virion transfer to the elusive secondary receptor is slow, and this may be for several reasons: the affinity for the secondary receptor may be low, there may be few secondary receptors available for binding on the cell surface and/or there are few binding sites available on the HPV capsid (90, 105). Several cellular molecules are proposed as either entry or co-receptors for HPV internalisation such as α-integrins, tetraspanins, epithelial growth factor receptor (EGFR), keratinocyte growth factor receptor (KGFR), and annexin A-2. These host molecules may act co-operatively to enable HPV infection, as no single receptor was sufficient for HPV infection (36, 53, 82, 108, 134, 135).

Integrins recognize short peptide sequences present on many ECM proteins, and binding leads to activation and intracellular signalling (101). HPV16 VLP binding was found to correlate with α-integrin expression, with HaCaT keratinocyte and cervical cancer cells having the greatest integrin expression and HPV16 binding, while non-keratinocytes had low levels of both integrin expression and HPV binding. Therefore, integrins may be a factor in the restricted cell tropism demonstrated by HPV. Treatment of HaCaT cells with an α6 integrin monoclonal antibody reduced HPV6b VLP binding, and HPV types 6b, 18, 31, 35 and BPV1 were found to induce intracellular signalling upon binding to integrins. These studies indicate a direct interaction between integrins and HPV (136, 137). However, increased β integrin expression on cells did not correlate with HPV16 VLP binding, and antibodies directed against non α-integrin types did not inhibit HPV binding, suggesting α-integrin are the integrin type used during HPV infection (136, 138). HPV may use the L1 protein to bind α-integrin as both L1

only and L1/L2 VLPs could induce signaling (137). Interestingly, cell surface-bound HSPGs co-operate with integrins, allowing cell attachment to the ECM, cell-cell interactions and cell motility (94). Potentially, HPV uses this interaction to form a receptor complex with cell surface HSPGs for entry into the cell, which could explain why no single secondary receptor has yet been identified (108). In contrast, other work has demonstrated that HPV entry may be integrin independent, suggesting that other cellular factors are more important for HPV entry (79, 96, 100).

Keratinocytes infected with HPV16 virions displayed colocalization with tetraspanins CD63 and CD151 on the cell surface of HeLa cells, implying that tetraspanins may be an HPV cell surface receptor. However, CD151 siRNA knockdown did not diminish HPV16 PsV binding to HeLa, HaCaT or primary human keratinocytes but did inhibit HPV16 PsV infectivity. This indicates that tetraspanins may function during intracellular steps in HPV infection. Likewise, only HPV PsVs associated with CD151 were found intracellularly (139, 140). Similarly, HPV may use growth factors and their receptors for internalisation, as demonstrated for EGFR and KGFR on the plasma membrane of HaCaT cells, leading to rapid activation of the growth factor receptors and subsequent activation of their cell signalling pathways (108). No single growth factor receptor was proposed to be responsible for HPV entry. When EGFR and KGFR were added to serum-free medium, HPV16 infection of HaCaT cells was restored, further indicating that growth factors are important receptors for HPV during infection (108).

Another surface protein suggested to be involved in HPV infection is annexin-A2, which was found to be present on both HeLa and HaCaT cells. An annexin-A2 antibody reduced HPV16 PsV binding to both keratinocyte cell lines, while shRNA knockdown of annexin-A2 reduced HPV16, 18, 31 and 45 PsV binding to HeLa cells. Moreover, re-expression of annexin-A2 in these cells restored HPV16 PsV infectivity. Co-IP confirmed an interaction between HPV16 PsVs and annexin-A2, which may be mediated by the HPV L2 protein, as the L2 amino acids from 108 to 126 bound annexin-A2, suggesting L2 may be functional in HPV entry (141, 142).

Given the above evidence, it is well established that, after initial HSPG binding, HPV is prepared for entry and intracellular infection steps by several host factors including cyclophilin B, kallikrein 8 and furin, before transfer onto a single or several secondary entry receptor(s).

1.6 HPV internalisation and trafficking

HPV internalisation into host cells is a slow process and is presently thought to occur by a process similar to macropinocytosis, as the process is dependent on actin dynamics, Na+ /H+

exchangers, P21 activated kinases (PAK-1) and protein kinase C (PKC) signalling (59, 89, 134, 143, 144). Interestingly, actin is also necessary for internalisation of cell surface HSPGs, the initial HPV receptor. Potentially, HPV also makes use of this pathway to enter the cell (145). This idea was confirmed when HPV was internalised into a non-infectious cellular pathway in the presence of the actin-depolymerizing drug cytochalasin D, suggesting that intact actin filaments are required in early events during HPV entry (92). Further, HPV31 virions were found to cause cytoskeletal rearrangements and filopodia formation on HaCaT cells which was activated by tyrosine kinase and phosphoinositide 3 (PI3)-kinase signalling pathways (146). HPV31 infection of HaCaT cells was inhibited by inhibitors of tyrosine kinases, PI3 kinases, actin polymerization, and microtubule polymerization and the viruses remained on the cell surface. In contrast, during HPV31 infection of untreated cells, the viruses were detected in the nuclear region. This reinforces the notion that these cell processes are important for HPV internalisation and trafficking (134). However, HPV internalisation is independent of clathrin, caveolin, cholesterol and dynamin (105, 134, 143, 144).

After internalisation, HPV is trafficked through the cell cytosol in endosomal compartments, and the acidification of late endosomes allows viral capsid uncoating. This process may require annexin-A2. L2 complexed with the HPV genome escapes the endosome and travels to the nucleus via the trans-Golgi network (113, 142-145). Endosomal escape is reliant on successful furin cleavage of L2 and requires several cellular proteins including γ-secretase and sorting nexin 17 (144) (147-149). Cyclophilins within the host cell facilitate detachment of the L2 capsid protein with the viral genome from the HPV capsid during capsid disassembly (150). Following this, the L2 viral genome complex travels through perinuclear compartments, which requires an interaction between L2 and the dynein microtubule network (143, 145, 151, 152). The process of nuclear entry by the viral genome is unclear, although it does involve L2 and the viral genome co-localising with nuclear import proteins, nuclear domain 10 (ND10) or promyelocytic leukaemia (PML) bodies, with nuclear envelope breakdown (105, 144, 153, 154). Cell cycle progression may be necessary for successful HPV infection (75). This is likely as the infectious cycle of HPV is closely linked to that of its' host cell. A summary of the entry mechanisms used by HPV16 to enter host cells is presented in Figure 1.6.

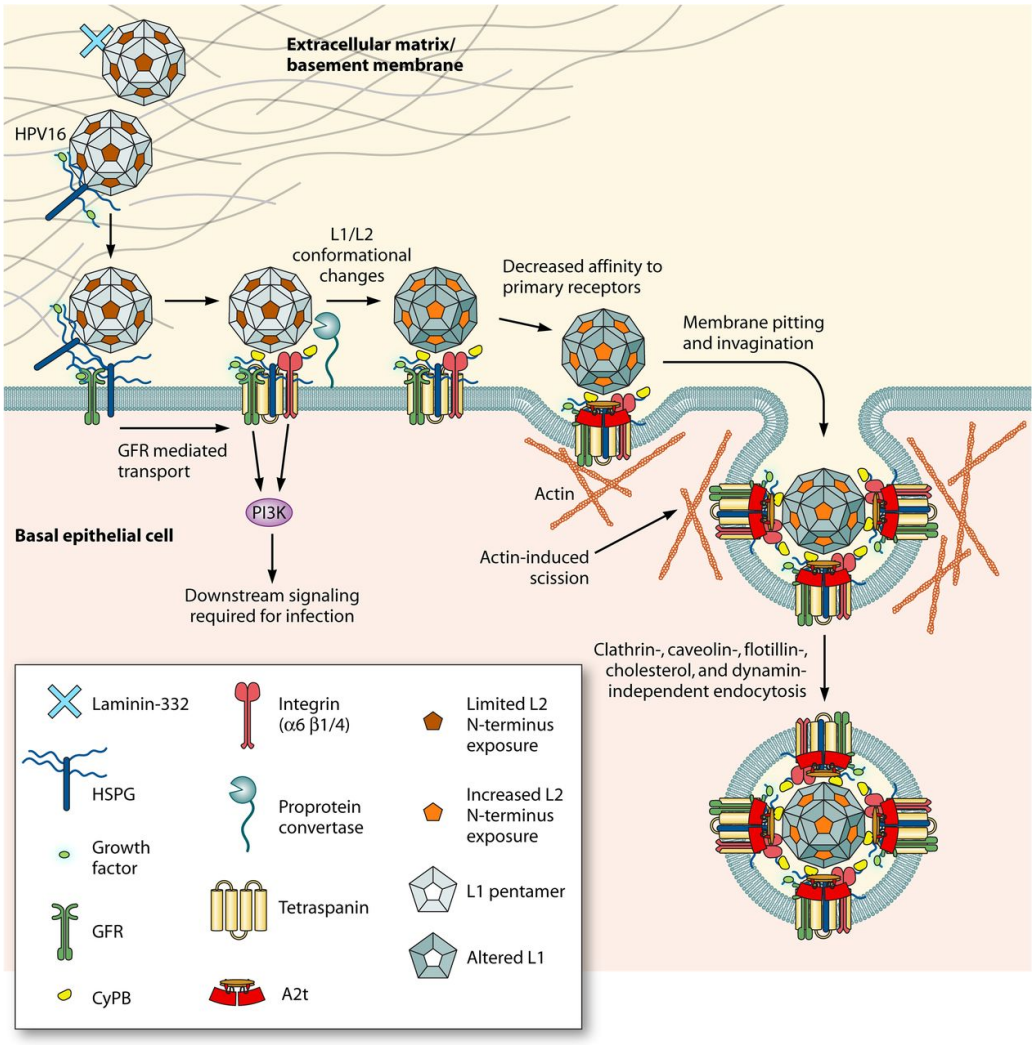


Figure 1.6. Schematic of entry mechanism used by HPV16 for successful infection. HPV16 binds to heparan sulphate proteoglycans (HSPGs) present on the extracellular matrix or cell surface. Epidermal growth factor and keratinocyte growth factor may also be activated upon cell surface binding and begin intracellular signalling cascades for virus internalisation. After HPV binding to the cell surface cylophillin B may cause a conformation change in the capsid protein, exposing the N-terminus of L2. HPV16 may also bind to α6-intergrin, initiating an intracellular singalling cascade. HPV16 then transfers to the secondary receptor (complex) and is internalised via clathrin- caveolin, lipid raft-, flotillin-and dynamin-independent endocytosis. Figure adapted from (53)

1.7 Challenges of studying HPV infection in the laboratory

As papillomaviruses are difficult to propagate in monolayer cell culture, *in vitro* studies are problematic without the use of other papillomavirus production systems, such as making use of pseudovirions (5, 6, 85). HPV pseudovirions (HPV PsVs) are HPV based gene transfer vectors, which are mostly indistinguishable from native HPV virions and have been successfully used to study HPV antigenicity, receptor usage, entry mechanisms and capsid structure (36, 50, 105, 126, 155, 156). They are essentially viral particles which are made up of the L1 and L2 structural proteins, forming the ~55 nm capsid encapsulating a reporter plasmid, a “pseudogenome” (85, 126, 157). The pseudovirion capsids are structurally similar to the native virions, composed of the correct disulphide bonds linking capsomers, due to the maturation step during PsV production (section 2.2) (56, 105, 157). This *in vitro* system is also flexible, as no specific packaging sequence is required, and allows high titres of HPV particles to be produced (85, 105, 126). For these reasons, this is the HPV production used in this investigation.

1.8 Hypothesis

While much is known about the HPV lifecycle and the virus' potential to cause cancer (section 1.4), and preventative measures like vaccines exist and are safe (section 1.2), there is much that is inconclusive regarding HPV entry and infection of host cells. Clarifying HPV entry into cells could aid in the development of inhibitors, to prevent HPV infection and possible cancer development.

In line with the proposed indirect evidence (52, 105, 106, 128, 158), this study hypothesised that furin cleavage of the HPV16 minor capsid protein L2 leads to changes in the quaternary capsid structure which can be visualised by high-resolution EM techniques. These changes are hypothesised to have consequences for reducing the affinity of HPV16 PsVs for the initial cell surface receptor, HSPGs, and allowing virus transfer to the entry receptor or receptor complex.

1.9 Aims and Objectives

Aim 1: To characterise the effects of furin cleavage on HPV16 PsVs through *in vitro* incubation of the viral particles with exogenous furin, high-resolution imaging using transmission electron microscopy (TEM) and single particle analysis.

This aim will be addressed by the following objectives:

Objective 1.1: To determine whether the addition of exogenous furin can cleave the L2 protein within the HPV16 PsV capsid. This will be assessed by SDS-PAGE and silver staining, and the results will be compared to furin uncleaved HPV16 PsVs. A cell infection assay will also be conducted to confirm the effect of furin cleavage on HPV16 PsV infectivity.

Objective 1.2: To establish whether furin cleavage of L2 within HPV16 PsVs affects the overall capsid morphology, which will be assessed by cryo-electron microscopy and single particle analysis of furin cleaved HPV16 PsVs in comparison to uncleaved virions.

Aim 2: To further characterise the effects of furin cleavage on the HPV capsid using computer simulated molecular dynamics.

This aim will be addressed by the following objectives:

Objective 2.1: Determine whether any conformational changes seen in the furin cleaved HPV16 PsVs affect the ability of the HPV16 PsV to bind to an anionic peptide as a substitute for heparin. This will be assessed using molecular dynamics software and generated models of HPV16 and the anionic peptide.

2 Methods

2.1 Materials

Unless otherwise stated, chemicals used during this project were of analytical or reagent grade. Cell culture media, SDS-PAGE gels and other solutions made up in the laboratory are detailed in the Appendix section. Reagents were generally purchased from Merck or Sigma-Aldrich, unless indicated differently.

2.2 HPV16 PsV production

2.2.1 Using HPV16 PsVs to study HPV infection

To study HPV16 infection *in vitro*, infectious HPV16 pseudovirions (HPV16 PsVs) were produced according to established protocols (159-161). Papillomavirus pseudovirions can encapsulate some cellular DNA during production and when making use of human-derived virus packaging lines, such as human embryonic kidney 293TT (HEK293TT) cells, oncogenes like the SV40 large and small T antigens present in these cells and resulting pseudovirions could be a safety risk (85). Therefore, pseudovirion production was done under biosafety level two (BSL-2) conditions.

HPV16 PsVs are essentially gene transfer vectors and are made of two mammalian expression plasmids: pXULL and pCMV GLuc 2 (New England Biolabs) (126, 160). pXULL is a codon optimised HPV16 L1 and L2 capsid protein expression plasmid, with an SV40 origin of replication (ori). The SV40 ori allows for increased production of the two HPV capsid proteins in cells that also express high levels of the SV40 large T antigen, such as HEK293TT cells (159, 160). pCMV GLuc 2 is a reporter plasmid representing the HPV pseudo-genome, containing a cytomegalovirus immediate early gene promoter (pCMV) and a Gaussia luciferase (GLuc) reporter gene, allowing quantification of successful cell infection once the reporter plasmid has reached the infected cells' nucleus. Luciferase activity can be easily measured as a read-out for infection. Gaussia luciferase is a ~19 kDa bioluminescent enzyme derived from *Gaussia princeps* and is secreted into the cell culture medium upon successful HPV-PsVs infection (159-162).

Siliconized 1.5 mL and 2 mL Eppendorf tubes were used to prevent loss of HPV16 PsVs, as these PsVs may adhere to plastic (162, 163).

2.2.2 Cell culture

HEK293TT cells, the immortalised human keratinocyte cell line HaCaT and the immortalised human epithelioid cervical carcinoma cell line HeLa (ATCC® CCL-2™) were cultured in Dulbecco's Modified Eagle Medium (DMEM; Sigma Aldrich), supplemented with 10% Foetal Calf Serum (Biochrom) and the antibiotics penicillin (100 U/mL) and streptomycin (100 U/mL) (159, 160). Cells were cultured in 10cm cell culture dishes (SPL Life Sciences) and incubated in a humidified atmosphere with 5% CO_2 at 37°C. All cell lines have been authenticated by IDEXX BioAnalytics and all cell culture procedures were done under sterile conditions to decrease the risk of contamination. Additionally, mycoplasma testing of cultured cells was done routinely, to assess for mycoplasma infection.

For the antibody neutralisation HaCaT cells were seeded in 12-well plates, at approximately $5x10^4$ cells per well. The following day, cells were infected with HPV16 PsV-antibody complexes (section 2.5.3) (126, 164).

2.2.3 Plasmid preparation

pXULL and pCMV-GLuc 2 (New England Biolabs) were amplified in *E. coli* and grown in selection media containing the antibiotics ampicillin (75 µg/mL) for pCMV-GLuc-2 and zeocin (50 µg/mL) for pXULL. Plasmids were purified using the PureYield™ Plasmid Maxiprep system (Promega).

2.2.4 Production of HPV16 PsVs

For HPV16 PsV production, HEK293TT cells were seeded onto thirty 10 cm cell culture dishes (Greiner Bio-One) in complete cell culture medium and grown overnight. After this, the calcium phosphate transfection method was used to transfect the HEK293TT cells (126, 160). Each 10 cm cell culture dish was transfected with each of the following plasmids: 5 µg pXULL plasmid and 12 µg pCMV-GLuc 2 with 25 µl $CaCl_2$ (final concentration 0.625 M) and 200 µL 2xHBS buffer (Appendix). After 48 hours, HEK293TT cells were harvested by trypsinisation and centrifuged in a 50 mL Falcon tube at 2000 rpm for 2 minutes. The cell pellet was resuspended in 5 mL 1xphosphate buffered saline (PBS; Appendix) and centrifuged again. The resulting cell pellet was resuspended in 1 mL 1xPBS and transferred to 2 mL siliconized Eppendorf tubes. The cell pellet was centrifuged at 3000 rpm for 3 minutes, after which the PBS was removed. Then the cell pellet was resuspended in the same volume of PBS, and the following was added: 9.5 mM $MgCl_2$, 0.35% Brij®-58 (Polyethylene glycol hexadecyl ether, Sigma Aldrich) for solubilizing proteins, 2 µL Benzonase® (Sigma-Aldrich) and 2 µL Exonuclease V, to digest un-encapsulated DNA.

The cell lysate was incubated at 37°C for 24 hours, with occasional mixing, to allow for HPV16 PsV inter-L1 disulphide bond formation and capsid maturation (56, 89, 160, 165). After pseudovirion maturation, the lysate was chilled on ice for 5 minutes before 0.17 volumes of 5 M NaCl in 1xPBS was added. This was further chilled on ice for 10 minutes and stored at -80°C. The lysate then underwent three freeze and thaw cycles at -80°C and 25°C.

2.2.5 Purification of HPV16 PsVs

Separation of pseudovirion capsids from cell debris and detergent is done by a high salt extraction step, followed by ultracentrifugation (56, 160, 161). The cell lysate was centrifuged at 8000g for 10 minutes at 4°C, the supernatant was collected in siliconized Eppendorf tubes, while the remaining cell pellet was washed with 400 µL High Salt Buffer (HSB) (Appendix) and centrifuged at 8000 g for 10 minutes at 4°C. Thereafter, the supernatants were combined, and the volume was brought up to 3.5 mL with 1xHSB.

The cleared cell lysate was then loaded onto a discontinuous caesium chloride (CsCl) gradient to further purify the HPV16 PsVs from the cell lysate. This CsCl gradient was prepared as follows: 4 mL light CsCl (1.25 g/mL in 1 mL 1xHSB was pipetted into a 13.2 mL Ultra-Clear™ centrifuge tube (Beckman Coulter, product #344059). A layer of 4mL heavy CsCl (1.4 g/mL in 1 mL 1xHSB (Appendix) was gently pipetted under the light CsCl layer and the interface between the heavy and light layers was marked on the tube. Finally, the cell lysate was gently pipetted onto the light CsCl layer. The sample was centrifuged for 16-17 hours at 20 000 rpm at 4°C in a swinging bucket rotor (Beckman SW40Ti). Following ultracentrifugation, the centrifuge tubes were removed and clamped with a ring-stand apparatus within a biosafety cabinet. HPV16 PsV were then extracted with an 18-gauge needle attached to a 5 mL syringe, from the faint band slightly above the CsCl interface. The remaining liquid within the Ultra-Clear™ centrifuge tube was discarded into a beaker containing Virkon™ (Vetoquinol) disinfectant.

The harvested HPV16 PsV samples were transferred to an Amicon® Ultra-4 filter device with 100 000 kDa molecular weight cut off (Milipore #UFC810024) and centrifuged at 3 000 rpm for 10 minutes. The samples were washed with 3 mL 1xHSB buffer and centrifuged at 3 000 rpm for 10 minutes again. This was done to remove any remaining CsCl and concentrate the HPV16 PsV sample. The HPV16 PsV preparations were then pipetted into siliconized Eppendorf tubes and stored at - 80°C (160).

2.3 Bicinchoninic assay

Protein concentration of each HPV16 PsVs preparation was measured using the Pierce™ BCA Protein Assay Kit (Thermo Scientific), according to the manufacturer's instructions.

2.4 Furin cleavage of HPV16 PsVs

Furin cleavage of HPV16 PsVs was performed according to Schäfer *et al*. (126). Briefly, purified HPV16 PsVs (~0.35 mg/mL) were incubated with furin (1.8 U/µg; New England Biosystems) in siliconized Eppendorf tubes at 37°C for 14 hrs (126). Successful furin cleavage was assessed by SDS-PAGE (section 2.5.1) followed by silver staining (section 2.5.2) and further supported by a cell culture infection assay (section 2.5.4).

2.5 Assessment of HPV16 pseudovirion preparation quality

2.5.1 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) is used to separate proteins based on their molecular weights and is used as an assessment for protein purity and a qualitative check on the amount of protein present within a sample. Therefore, this technique was used to assess purity of HPV16 PsV preparations as conducted previously (47, 126).

SDS-PAGE gels consisted of 3% stacking gel and either 7.5% or 10% separating gel (Appendix) as indicated. 5X Pierce™ Lane Marker Reducing Sample Buffer (Thermo Scientific) containing Dithiotreitol (DTT) was added to experimental HPV16 PsV samples (0.1 µg/µL) which were heated at 95°C for 10 minutes, for protein denaturation. The denatured pseudovirion samples were then loaded onto the set SDS-PAGE gels and protein samples were electrophoresed in 1X running buffer (Appendix), according to established protocols (47, 126). A protein molecular weight marker, Colour Pre-stained Protein Standard, Broad Range Marker (11-245 kDa; New England Biolabs) was used to estimate loaded protein sample molecular weights. Electrophoresed gels were stained for protein visualisation and assessment (section 2.5.2).

2.5.2 Pierce™ silver staining

SDS-PAGE gels were stained to visualise the HPV16 PsV L1 and L2 proteins (and any possible contaminants) using the Pierce™ silver stain kit (Thermo Scientific), according to

the manufacturer's instructions. Briefly, SDS-PAGE gels were washed twice for 5 minutes in ultrapure dH_2O. Gels were then fixed twice for 15 minutes each, using a 30% ethanol:10% acetic acid solution. Afterwards, the gels were washed twice for 5 minutes each with 10% ethanol, followed by two washes in ultrapure dH_2O. Gels were then sensitized for 1 minute with sensitiser solution made according to the protocol (Pierce™ silver stain kit; Appendix), followed by two 1 minute washes in ultrapure dH_2O. Gel staining was done for 30 minutes using the staining solution (Pierce silver stain kit; Appendix), followed by two quick 20 second washes with ultrapure dH_2O. Thereafter, gels were developed for 2 to 3 minutes development using the developer working solution (Pierce silver stain kit; Appendix) until protein bands appear. The developing reaction was then stopped with 10% acetic acid for 10 minutes. An HPV16 PsV L2 protein band was expected at 75 kDa and a prominent HPV16 PsV L1 protein band was expected to co-migrate at 55 kDa in pure HPV16 PsV preparations (47, 112).

2.5.3 Antibody neutralisation of HPV16 PsVs

To assess quality and infectivity of the pseudovirion preparations, HPV16 PsVs were pre-incubated with either the HPV16 neutralising antibody (H16.V5; kindly provided by Neil D. Christensen, Pennsylvania State University College of Medicine, USA) or the HPV18 neutralising antibody (HPV18.J4; kindly provided by Neil D. Christensen, Pennsylvania State University College of Medicine, USA) at a final concentration of 1:1000 for 1 hour at 4°C (126, 159, 164). HaCaT cells were seeded according to Schafer *et al*. (126, 164) in 75 cm^2 cell culture flasks (SPL Life Sciences). These antibody-complexed HPV16 PsVs were then added to HaCaT cells and resulting luciferase secretion by infected cells was measured as a read-out for HPV infection (section 2.5.3). H16.V5 was expected to neutralise (i.e. abolish) infection of HPV16 PsVs, thereby acting as a positive control, while the HPV18 antibody acted as a negative control, and was not expected to affect HPV16 PsV infection of HaCaT cells.

2.5.4 Cell infection assay with furin cleaved HPV16 PsVs

To assess the infectivity of HPV16 PsV preparations, cell infection assays were conducted using HeLa cells, which were seeded in monolayer culture as described (section 2.2.2) (126). Furin cleaved or furin uncleaved HPV16 PsVs were added to cultured HeLa cells and luciferase secretion into the cell supernatant was measured 48 hours later (section 2.5.5).

2.5.5 Luciferase assay

Luciferase assays were performed to quantify the infectivity of HPV16 PsV preparations in HeLa and HaCaT cells, where the infection was measured 48 hours after infection. 5 µL cell supernatant was added to a white 96-well plate (Nunc) and luciferase activity was measured using the Gaussia luciferase (GLuc) Assay Kit (New England Biolabs) together with a GloMax® Explorer Multimode Microplate Reader (Promega). Per sample, freshly prepared GLuc Assay Solution made up of 0.5 µL BioLux GLuc Substrate with 50 µL BioLux Assay Buffer (New England Biolabs) was added per well. Luciferase luminescence was measured with 2 seconds lag time over 10 seconds integration time.

2.6 Increasing the concentration of furin cleaved HPV16 PsVs

For imaging by cryo-electron microscopy, furin cleaved HPV16 PsVs were further concentrated. A small steel rod (61.57 mm^3) was added to the bottom of an Amicon® Ultra-4 filter device with 100 000 kDa molecular weight cut off (Milipore #UFC810024) to minimise the loss of HPV16 PsVs within the filter. 38 µL of the furin cleaved HPV16 PsV preparation (~0.35 mg/mL) was added to the filter, followed by 250 µL 1xHSB. This was then centrifuged at 14 000x g at 4°C in a benchtop centrifuge for 20 minutes. 250 µL 1xHSB was added and the centrifugation step was repeated twice more, followed by three rounds of 15 minute centrifugation steps. In each case, 250 µL 1xHSB was added. The bottom of the Amicon® filter was emptied and the sample was spun for another 15 minutes without adding HSB. Finally, 13 µL HSB was added to the concentrated HPV16 PsV, making the concentration ~0.6 mg/mL, before transferring the sample to a siliconized Eppendorf tube and preparation of the cryo-electron microscopy (cryo-EM) grids (section 2.10).

2.7 Negative stain Electron Microscopy

Negative stain transmission electron microscopy (TEM) was used to visualise HPV16 PsVs, to assess the overall structure, dispersion, and homogeneity of the preparations.

2.7.1 Grid preparation

Carbon coated copper grids (Electron Microscopy Sciences) were glow-discharged in air at a plasma current of 25 mA for 30 seconds using an EMS100× Glow Discharge Unit (Electron Microscopy Sciences) to render the grids more hydrophilic and negatively charged. 2.5 µL of HPV16 PsV preparation was pipetted onto the glow-discharged grid and allowed to adhere for

30 seconds. The excess sample was then gently and indirectly blotted off each grid with filter paper. Each grid was then washed with 3 drops of distilled water (dH_2O) followed by staining with 2 drops of 2% aqueous uranyl acetate (SPI supplies), indirectly blotting the grid dry with filter paper wedges between each step. Each grid was then air dried before transmission electron microscopy (TEM).

For the furin uncleaved HPV16 PsVs, processing by furin and increasing the concentration of the virions was not done (section 2.6). The furin unprocessed HPV16 PsVs (~0.4 mg/mL) were directly added to the prepared grids after production and storage at -80°C (section 2.10).

2.7.2 Transmission electron microscopy

Each negative stain grid was loaded, using a single side-entry grid holder (Gatan), onto a Tecnai T20 Transmission Electron Microscope (FEI) with an LaB_6 electron gun operating at 200kV, and equipped with a CCD camera (2048 x 2048-pixel; Gatan). Micrographs were collected at SA 53 000 x magnification in low dose mode.

2.8 Cryo-EM grid preparation for F20 TEM

Quantifoil® 'holey' R2/2 carbon coated copper grids were glow-discharged in air at a plasma current of 25 mA for 30 seconds using an EMS100× Glow Discharge Unit (Electron Microscopy Sciences) to render the grid more hydrophilic and negatively charged. The concentration of HPV16 PsVs (~0.5 mg/mL) was sub-optimal for cryo-EM, therefore carbon coated cryo-EM grids were used to increase particle adsorption to the grid and the number of particles available for imaging. A Vitrobot Mark IV (Thermo Scientific) was used to apply, blot, and vitrify each sample. 3 µL of HPV16 PsV was pipetted onto glow discharged grids and allowed to adhere in a 100% humidified chamber for 30 seconds, before being blotted with filter paper and plunged into a bath of liquid ethane, cooled by liquid nitrogen. The vitrified grid was then transferred into a labelled grid box and this was placed inside a 50 mL Falcon.

2.9 Transmission electron microscopy and data collection

The vitrified cryo-EM grid was loaded onto the single grid side-entry cryo-holder (Gatan) under liquid nitrogen. The specimen holder was then inserted into the Tecnai F20 microscope (Phillips/FEI) and allowed to equilibrate to about -180°C for 90 minutes before data collection. Micrographs were collected using the parameters listed in Table 3.2.

2.9.1 Micrograph pre-processing

The micrographs were stored locally for processing on a computer with 128Gb RAM and four 1070 Nvidia GPUs. These micrographs were imported into REgularised Likelihood OptimisatioN (RELION) 2.1 software for three-dimensional reconstruction (166-170). Contrast transfer function (CTF) estimation was performed using Gctf software, with dose-weighting (171). Particles within the micrographs were manually picked from micrographs, and normalised. Then, the contrast was inverted, and the particles were extracted in a box size of 600 pixels without rescaling.

2.9.2 Three-dimensional reconstruction

Two-dimensional (2D) class averages of the extracted particles were calculated which are compiled of several aligned particles with similar features for a particle projection or orientation (172). The particles were averaged into ten 2D classes in twenty iterations, with a regularisation parameter T = 2. 2D classes had a mask diameter of 600 Å and the images were aligned. There was an in-plane angular sampling of 6° with an offset search range of 5 pixels and an offset search step of 1 pixel. The pooled particles were pre-read into RAM with four GPUs used, split over four MPI processes and two threads. A subset of classes was selected from the 2D class averages generated after re-centering. These classes were selected as they represented particles with the highest resolution in the data collected and underwent three-dimensional (3D) classification. The 3D classification was done on the selected 2D classes, with a featureless sphere generated in the software System for Processing Image Data from Electron microscopy and Related fields (SPIDER) as the reference map, which was lowpass filtered to 60 Å (173). Icosahedral (I2) symmetry and regularisation parameter T = 4 was applied. The mask diameter was 600 Å with an angular sampling interval of 3.7°, an offset search range of 5 pixels and an offset search step of 1 pixel. Resulting 3D classes were visualised using UCSF Chimera ('Chimera') for analysis (174). The 3D classes with the highest resolution were selected and used for 3D refinement. The best 3D class was used at 60 Å lowpass filtered reference map and icosahedral (I2) symmetry was applied. Initial angular sampling was 7.5° as HPV is an icosahedral virus, with a local search from auto sampling 1.8°, initial offset search range of 5 pixels and an offset search step of 1 pixel, which would be increased as the 25 iterations proceeded. The resulting 3D refinement model was assessed using UCSF Chimera (section 2.14) and the Fourier Shell Correlation (FSC) curve.

2.10 Cryo-EM grid preparation for high resolution imaging

Quantifoil® 'Holey' R2/2 carbon-coated copper grids (Electron Microscopy Sciences) were used for both the furin cleaved and furin uncleaved HPV16 PsV samples. An additional thin layer of carbon was deposited onto the holes within the copper grids for the furin cleaved HPV16 PsV samples. This was not done for the furin uncleaved HPV16 PsV samples. 'Holey' R2/2 carbon coated copper grids (Quantifoil) were covered with an additional layer of thin amorphous carbon before being glow-discharged in air at a plasma current of 25 mA for 30 seconds using a EMS100× Glow Discharge Unit (Electron Microscopy Sciences, USA) to render the grid more hydrophilic and negatively charged.

Sample application to cryo-EM grids and vitrification was done using a Vitrobot (Thermo Scientific). The concentration of HPV16 PsVs (~0.6 mg/mL) was sub-optimal for cryo-EM, therefore carbon-coated cryo-EM grids were used to increase particle adsorption to the grid and the number of particles available for imaging. 2.5 µL of the HPV16 PSV preparations was pipetted onto the glow discharged thin carbon coated 'Holey' R2/2 carbon grids (Quantifoil) within the 100% humified chamber of the Vitrobot set at 4°C and allowed to adhere for 5 minutes. Subsequently, the excess sample was blotted away from both sides of the grid, for 30 seconds using the Vitrobot. The grid was then plunged into a bath of liquid ethane, cooled by liquid nitrogen, before being transferred to a labelled grid box under liquid nitrogen for storage. The grid boxes were then stored under liquid nitrogen within 50mL Falcon tubes, before being shipped to the Electronic Bio-imaging Centre (eBIC) at Diamond Light Source (UK).

2.11 High resolution transmission electron microscopy of HPV16 PsVs

The furin cleaved and uncleaved datasets were collected on G3 Titan Krios cryo-electron microscopes (Thermo Fisher Scientific), operating at 300kV with an X-FEG electron source (Schottky) at eBIC (Diamond Light Source, UK) equipped with either a K2 detector (Gatan; furin cleaved dataset) or a K3 detector (Gatan; furin uncleaved dataset) in counting mode with data acquisition using EPU software. Acquisition parameters are listed in Table 3.2 and Table 3.3. Each dataset was collected over several grid squares within a single cryo-EM grid, without major changes to microscope parameters, allowing the multi-frame movies collected to be combined into a single dataset.

Movies were collected in low-dose mode at a range of defocus values, to fill in the zeros of the contrast transfer function (CTF) although the defocus values were close to focus, to retain

as much high-resolution information as possible (172). The HPV16 PsV is a large and spherical proteinaceous virion, making it easily identifiable even close to focus, therefore it was unnecessary to use large defocus values (48, 172). Gain correction was done on both datasets prior to data collection, to normalise the response of each pixel within the detector.

2.12 Single particle analysis of furin uncleaved HPV16 PsVs

2.12.1 Movie pre-processing

All processing was done in the RELION 3.1 pipeline with wrappers to the relevant programs (MotionCor2, ctffind4), unless otherwise indicated (168, 170, 171, 175). The micrographs were transferred from Diamond Light Source to the local laboratory and stored locally, for processing. A summary of the single particle analysis for the furin uncleaved HPV16 PsVs is presented in Figure 2.1.

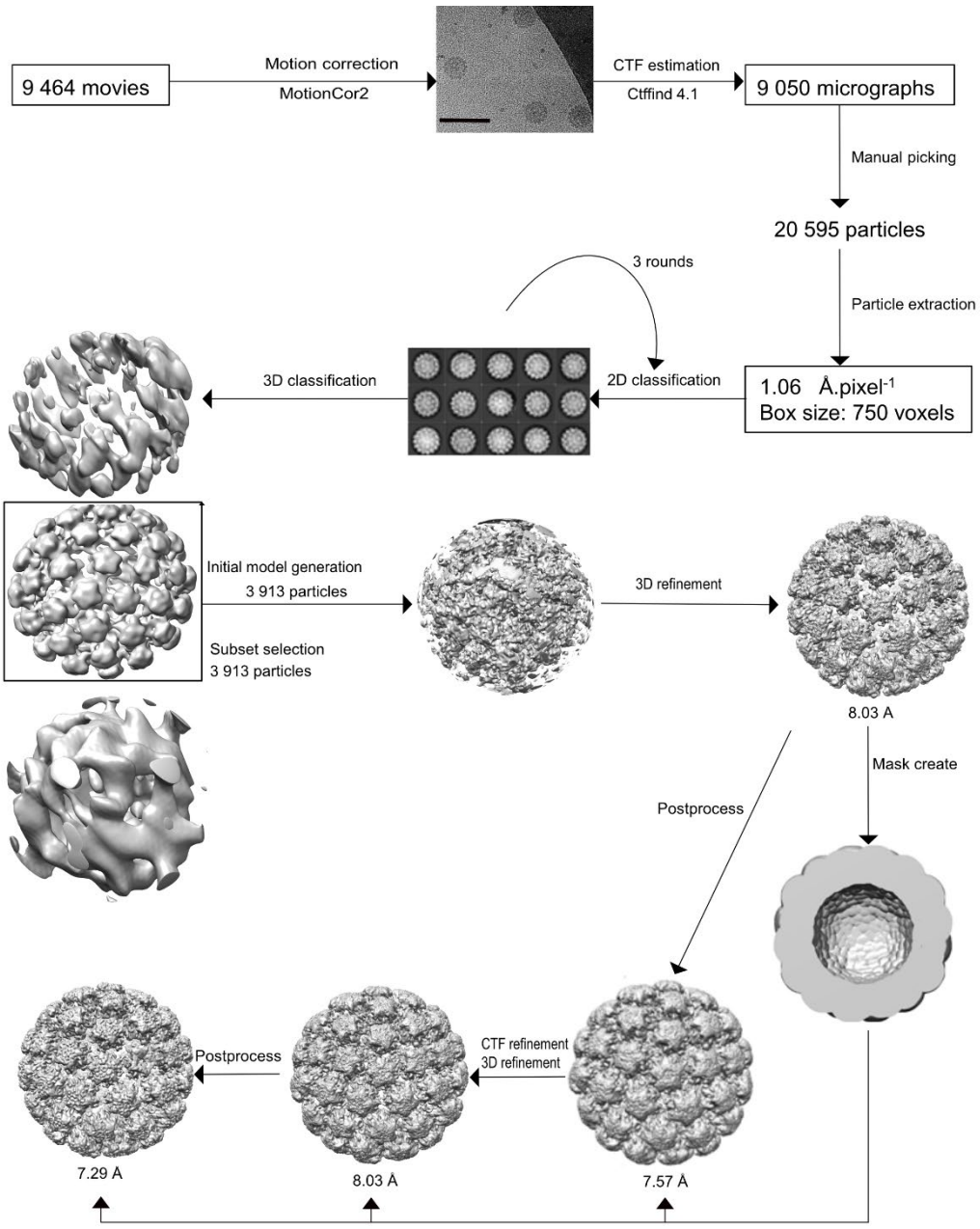


Figure 2.1, Single particle analysis of furin uncleaved HPV16 PsVs using RELION. Multi-frame movies were motion corrected using MotionCor2, and CTF parameters were estimated for each micrograph using Ctffind4. Individual HPV16 PsVs were extracted from the data using manually selected particle co-ordinates and subjected to three rounds of two-dimensional classification. After each round the best classes were chosen and the final particle set was used to generate a three-dimensional model which then underwent three-dimensional refinement. A solvent mask was then generated from this model with the interior map density deleted in UCSF Chimera. The particles were then subjected to three-dimensional classification followed by selection of the best class. Subsequently, the particles underwent three dimensional refinement and per particle defocus values were estimated during CTF refinement. Finally, the model was postprocessed with a β factor correction of -20.

Multi-frame movies of the furin uncleaved HPV16 PsV dataset were imported into the RELION 3.1 pipeline and MotionCor2 was used to correct for beam-induced motion during data collection (175), the micrographs were binned by a factor of 2, leading to scaling of 1.06 Å.pixel^{-1} as these movies were taken in “super-resolution” mode on a K3 detector (Gatan) with a pixel size of 0.53 Å/pix for 9 464 movies. Contrast Transfer Function (CTF) parameters were estimated using ctffind4 (176). CTF curves (with phase and amplitude terms) for corrected micrographs were assessed for fit to the theoretical CTF and micrographs with sub-optimal fits to the theoretical CTF were discarded, leaving 9 050 selected micrographs for further processing. Particle co-ordinates within micrographs were selected manually using RELION, producing a total of 20 595 particles from 8 862 micrographs. Particles were normalised, the contrast was inverted, and particles were extracted from micrographs in a box size of 750 pixels at a sampling of 1.06 Å. pixel^{-1}, with no downscaling. A small particle box size was used to minimise computational costs of downstream processing while retaining as much high-resolution data as possible.

2.12.2 Three-dimensional reconstruction of furin uncleaved HPV16 PsVs

2D class averages of the extracted particles were calculated to classify the extracted particles into 2D class averages based on projection image matching (170). Three successive rounds of reference free 2D classification were done to classify the extracted particles into fifty 2D class averages in twenty-five iterations, with a regularisation parameter T=2 and a solvent mask diameter of 800 Å. Particle images were aligned and an in plane angular sampling of 5 pixels, and offset search range of 5 pixels and an offset search step of 1 pixel was applied. After each round of 2D classification, the highest resolution class averages were selected, eliminating sub-optimal particles and random noise from the dataset. The particles from each class were also “regrouped” into ~ 200 classes in each 2D classification step, to improve estimates of their average noise power spectrum. This process resulted in 4 216 particles in 142 groups. Then, the particles were subjected to 3D classification with a previously generated furin uncleaved model as a reference map. The reference map was lowpass filtered to 60 Å, with no symmetry (C1) for 3D classification of the data into 4 classes. A regularisation parameter T = 4 was also applied to the twenty-five iterations, with a mask diameter of 700 Å. The particles producing the highest resolution 3D class average were selected, resulting in 3 913 particles in 124 groups.

An initial model was generated from the selected particles using RELION and the Stochastic Gradient Algorithm (SGD), with no symmetry (C1) applied (168, 177). A mask diameter of 700 Å was applied, with an initial angular sampling of 15°, an offset search range of 6 pixels, and offset search step of 2 pixels. The resulting model was 15 Å and was used as a reference map for an initial 3D refinement. The initial model was lowpass filtered to 60 Å, and icosahedral symmetry was applied (I2). A small 700 Å mask was applied and an initial angular sampling of 3.7 °, with a local search from auto sampling of 0.9°, as this was a large icosahedral virus. The initial offset range and initial offset step was 5 and 1 pixel(s) respectively. This resulted in a final resolution of 8.03 Å. The initial 3D refinement was lowpass filtered to 20 Å using relion_image_handler and was opened in the program UCSF Chimera. The map was segmented into regions and the density corresponding to genetic material inside the capsid was deleted, leaving a 'capsid-only' map (56). This map was saved and used in RELION to generate a solvent mask for downstream refinement. This solvent mask was lowpass filtered to 20 Å generated with an initial binarization threshold of 0.00103, the binary map was extended by twenty pixels and a soft edge of ten pixels was added. Postprocessing was done on the refined HPV16 PsV map, with the 'capsid-only' solvent mask. The post-processed map was sharpened with a user provided β factor of -20 and an *ad-hoc* lowpass filter of 8 Å. Per particle defocus values for the dataset were estimated. To further improve resolution, the data was subjected to a final 3D refinement and postprocessing as previously, with the 'capsid-only' solvent mask. The 3D refined map was sharpened with a user-provided β factor of -20 and an *ad-hoc* lowpass filter of 8 Å. This led to a final 3D reconstruction of the furin uncleaved HPV16 PsV at 7.29 Å, according to the "gold standard" FSC curve (167, 170, 178).

2.13 Single particle analysis of furin cleaved HPV16 PsVs

2.13.1 Movie pre-processing

All processing was done in the RELION 3.1 pipeline with wrappers to the relevant programs (MotionCor2, ctffind4), unless otherwise indicated (168, 170, 175, 176). The micrographs were transferred from Diamond Light Source to the local laboratory and stored for processing. A summary of the single particle analysis leading to a 3D reconstruction is presented in Figure 2.2. (167, 170, 178).

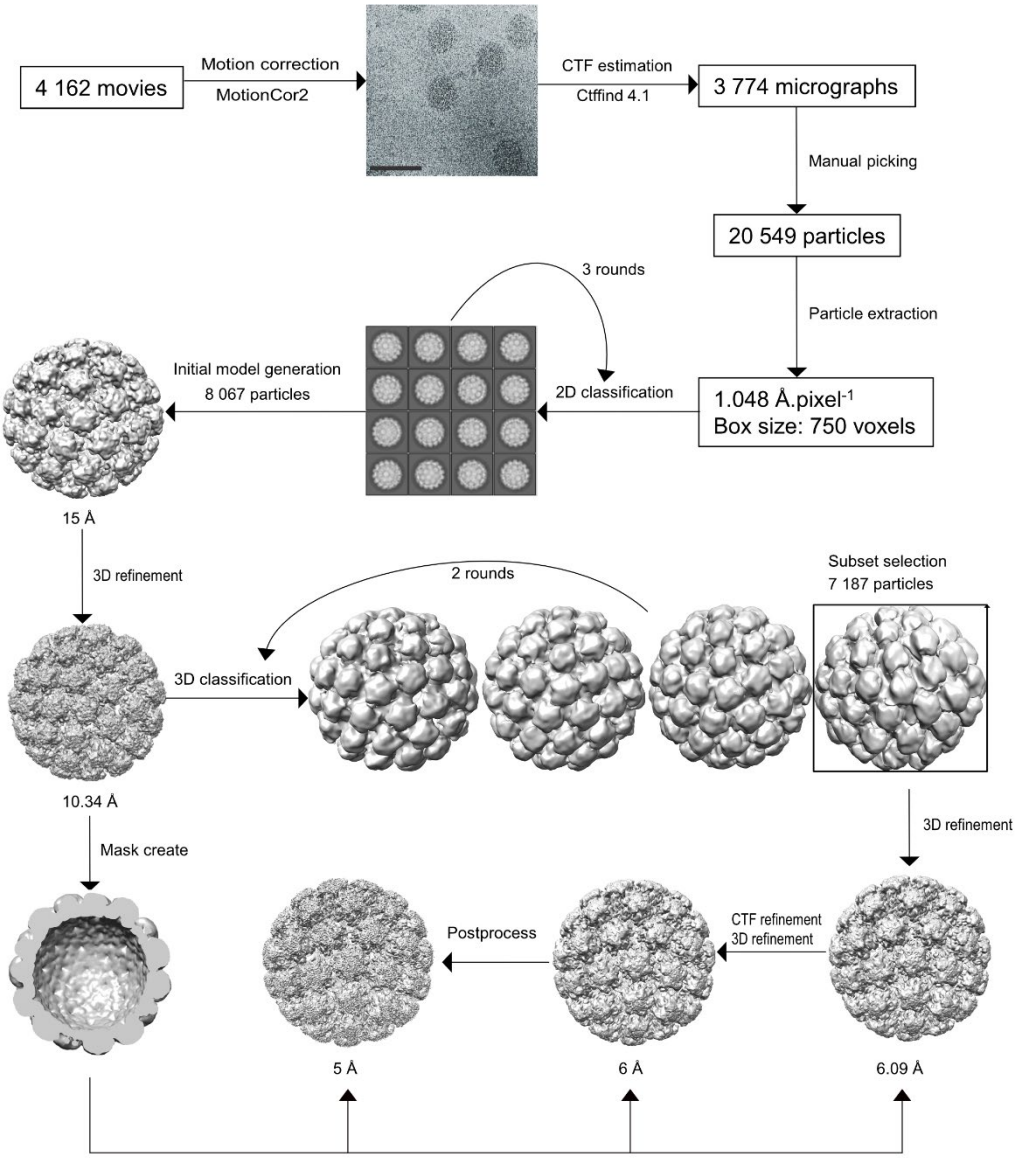


Figure 2.2. Single particle analysis of furin cleaved HPV16 PsVs using RELION. Multi-frame movies were motion corrected using MotionCor2, and CTF parameters were estimated for each micrograph using Ctffind4. Individual HPV16 PsVs were extracted from the data using manually selected particle co-ordinates and subjected to three rounds of two dimensional classification. After each round the best classes were chosen and the final particle set was used to generate a three dimensional model which then underwent three dimensional refinement. A solvent mask was then generated from this model with the interior map density deleted in UCSF Chimera. The particles were then subjected to two rounds of three dimensional classification with selection of the best classes after each round. Subsequently, the particles underwent three dimensional refinement and per particle defocus values were estimated during CTF refinement. Finally, the model was postprocessed with a β factor correction of -95.

Micrographs of the furin cleaved HPV16 PsV dataset were imported into the RELION pipeline, as the multi-frame movies had already been motion-corrected using MotionCor2, to correct for beam induced motion during data collection (175). These micrographs had a pixel size of 1.048 Å.pixel^{-1}, and data parameters in Table 3.3 were input into RELION. Contrast Transfer Function (CTF) parameters were estimated using ctffind4 (176). CTF curves (with phase and amplitude terms) for corrected micrographs were assessed for fit to the theoretical CTF and micrographs with sub-optimal fits to the theoretical CTF were discarded, leaving 3 774 selected micrographs for further processing. Particle co-ordinates within micrographs were selected automatically using RELION, and the resulting co-ordinates were curated manually, leaving a total of 20 549 particles from 3 551 micrographs. Particles were normalised, the contrast was inverted, and particle coordinates were extracted from micrographs in a box size of 750 pixels at a sampling of 1.048 Å.pixel^{-1}, with no downscaling.

2.13.2 Three-dimensional reconstruction

Three successive rounds of reference-free 2D classification were done to classify the extracted particles into 2D class averages based on projection image matching (170). After each round of 2D classification, the highest resolution class averages were selected, eliminating sub-optimal particles and random noise from the dataset. The particles from each class were also "regrouped" into ~ 500 classes in each 2D classification step, to improve estimates of their average noise power spectrum. This process resulted in 8 067 particles in 309 groups. Using RELION and the Stochastic Gradient Descent algorithm an initial model was generated from the selected particles, with icosahedral (I1) symmetry (168, 177). A mask diameter of 700 Å was applied, with an initial angular sampling of 15°, an offset search range of 6 pixels, and offset search step of 2 pixels. The resulting model was 15 Å and was used as a reference map for an initial 3D refinement. This high-resolution 3D refinement is an automated procedure, which calculates the Fourier Shell Correlation (FSC) for accurate determination of resolutions without over-fitting the data (168, 170). The 15 Å initial model was lowpass filtered to 60 Å, and no symmetry was applied (C1 symmetry). A 700 Å mask was applied and an initial angular sampling of 3.7°, with a local search from auto sampling of 0.9°, as this was a large icosahedral virus. The initial offset range and initial offset step was 5 and 1 pixel(s) respectively. This resulted in a model with a final resolution of 10.34 Å.

The initial 3D refinement was lowpass filtered to 20 Å using relion_image_handler and was opened in Chimera. The map was segmented into regions and the density corresponding to interior density inside the capsid was deleted, as this unordered region may affect the overall resolution of the finalised map, leaving a "capsid-only" map (56). This map was saved and

used in RELION to generate a solvent mask for downstream refinement. The solvent mask was lowpass filtered to 30 Å generated with an initial binarization threshold of 0.0044, the binary map was extended by five pixels and a soft edge of ten pixels was added. The map was also flipped 90° along the Z-axis of UCSF Chimera, to align with the data in RELION. The 10.34 Å map was lowpass filtered to 60 Å, and used a reference map, with no symmetry (C1) for 3D classification of the data into 4 classes. A regularisation parameter T = 4 was also applied to the twenty-five iterations, with a mask diameter of 750 Å. The particles producing the highest resolution 3D class average were selected, resulting in 8 067 particles grouped into 604 groups. A second 3D classification was done on these selected particles as previously, and the highest resolution 3D class averages were again selected, and 7 187 particles subjected to 3D refinement. 3D refinement was performed on the data, with a previously generated high-resolution HPV16 PsV map as the reference map, with the “capsid-only” solvent mask. The reference map was lowpass filtered to 30 Å and icosahedral (I2) symmetry, initial angular sampling of 3.7°, with a local search from auto sampling of 0.9° was applied. The initial offset range and initial offset step was again 5 and 1 pixel(s) respectively, leading to a reconstruction at 6.09 Å. Postprocessing was done on the refined HPV16 PsV map, with the “capsid-only” solvent mask generated. The post-processed map was sharpened with a user-provided β factor of -95 and an *ad hoc* lowpass filter of 5 Å. Afterwards, per-particle defocus values for the dataset were estimated. To further improve resolution, and the data was subjected to a final 3D refinement as in the previous refinement, and post-processing, with the “capsid-only” solvent mask. This led to a final 3D reconstruction of the furin cleaved HPV16 PsV at 5.006 Å, according to the “gold standard” FSC curve (section 2.14).

2.14 Resolution determination

The “gold standard” Fourier Shell Correlation (FSC) was used for automated resolution determination within RELION. When an FSC curve drops below 0.143, this is a good indicator of overall map resolution (178, 179). During 3D refinement particles are split into random halves and two 3D reconstructions are generated, which are independent of each other, to prevent ‘over-fitting’ of random noise in the data. This process also estimates the accuracy of angular assignments, allowing RELION to automatically determine when the 3D refinement has converged from the two half-reconstructions (168, 170). Comparison of each reconstruction with a published model to confirm the resolution achieved was also done (section 2.15).

2.15 Structure comparison

An independent and published HPV16 L1-only capsid asymmetric unit was retrieved from the Protein Data Bank (PDB ID: 5kep) and used as a model to confirm the handedness of the HPV16 PsV reconstructions produced here (50). UCSF Chimera ('Chimera') was used to view each HPV16 PsV reconstruction generated and to dock 72 copies (one for each HPV16 pentamer) of the L1 pentamer (part of the PDB 5kep structure) into the PsV reconstruction using the 'fitmap' command within Chimera, to optimise docking within the reconstruction (174).

Similarly, a high-resolution cryo-EM structure of an HPV16 quasivirion (EMB-6620) was retrieved from the Electron Microscopy Data Bank (EMDB) and docked into the reconstructions generated here, using Chimera, providing a qualitative measure of resolution obtained (50). As this quasivirion reconstruction was produced differently to our pseudovirion reconstructions, this structure was not used for evaluation of the effects of furin cleavage on the HPV capsid.

2.16 Assessment of HPV16 PsV conformation changes

The final furin cleaved (5 Å) and uncleaved (7.29 Å) HPV16 PsV reconstructions were assessed for any changes in conformation, which may be attributed to furin cleavage of L2. To achieve this, both models with the published cryo-EM structure of HPV16 (EMD-6620) were imported into Chimera. The two reconstructions generated in this project were docked into the published EMD-6620 map, using the 'fitmap' command in Chimera.

To generate atomic-level structures of our furin cleaved and uncleaved HPV16 reconstructions a published PDB of the L1-only capsid asymmetric unit was fitted into the two density maps. Each L1 chain within the HPV16 asymmetric unit (PDB ID: 5kep) was docked individually into either the furin cleaved or uncleaved HPV16 PsV reconstructions using 'fitmap' command in Chimera, known as rigid-body docking. Following this, the docked L1 chains of each reconstruction were combined into a single PDB, resulting in two PDB L1 models with each L1 subunit in either the furin cleaved or uncleaved conformation. The command matchmaker was used to align the L1 chain which forms part of the pentavalent capsomere (Chain A;

Figure 1.2[C]). Chain A was not aligned using ‘matchmaker’ in Figure 3.6 as this figure demonstrates that both hexavalent and pentavalent capsomeres move after furin cleavage of L2. The root mean squared deviation (RMSD) was used as a quantitative measure of differences between the generated furin cleaved and uncleaved structures, after alignment of chain A. RMSD was calculated by comparing all amino acids in each docked L1 chain (chain A to F) between the furin cleaved and uncleaved HPV16 PsV.

The distance that the L1 chains moved after furin cleavage of L2 around the canyon between hexavalent and pentavalent capsomere (the putative heparin-binding site) was measured in Chimera using the ‘distances tool’. Atoms within the L1 chains forming the asymmetric unit generated above were selected and the difference between the atom in the furin uncleaved versus the furin cleaved conformation was measured in angstroms (Å).

2.17 Molecular dynamics

Molecular dynamics was used to determine whether the change in capsid conformation observed between the furin cleaved and uncleaved HPV16 PsV reconstructions had any functional relevance for HPV16 attachment to heparin, a widely used substitute for HSPGs, the initial HPV cell attachment receptor (section 1.5).

After retrieving the two cryo-EM density maps from the EMDB a published structure of HPV16 (EMD-6620) was subtracted from the associated HPV16 with heparin bound structure (EMD-6619) using the ‘vop subtract’ command in Chimera, to generate a map containing heparin density, representing the putative HSPG binding site on the HPV16 capsid. This heparin density was lowpass filtered using to 8 Å using relion_image_handler (50, 167, 169, 170). An aspartic acid peptide was generated to simulate the length and negative charge of heparin, given the difficulties in generating molecular dynamics force fields for these highly heterogenous and flexible models. The peptide was generated in Chimera using the ‘Build Structure’ menu and is composed of eighteen repeating residues of aspartic acid, with alternating Φ = -57° and ψ = -47° for an α helix, or Φ = -139° and ψ = -135° for an anti-parallel β sheet, to generate a linear amino acid peptide. This peptide was rigid body docked into the heparin density map using the ‘fitmap’ command in Chimera. Following this, the aspartic acid peptide was combined with the three L1 chains (chains A, B and F) around the heparin density in either the furin cleaved or uncleaved conformation (section 2.16), using ‘copy/combine’ option in Chimera. This was done to generate two models; each composed of three HPV16 L1 chains, either in the furin cleaved or uncleaved conformation, with an aspartic acid peptide, to represent heparin, in the putative heparin location on the HPV16 capsid.

Finally, these models were imported into ChimeraX and a real-time molecular dynamics simulation was done using ISOLDE for both the furin cleaved and uncleaved HPV16 PsV models (180, 181). Hydrogens were added to the models and the simulation was run at 310 Kelvin overnight. The results were assessed manually and hydrogen bonds were measured between interacting amino acids on each L1 chain and the aspartic acid peptide, using the 'find H bond' option within ChimeraX, according to parameters specified in (182) and coloured black. The hydrogen bonding constraints were relaxed by 0.4 Å and 20°, and these hydrogen bonds were coloured orange.

2.18 Investigating the density maps for the location of L2

We investigated the location of the L2 protein within both the furin uncleaved and furin cleaved HPV16 PsV density maps, and whether after furin cleavage of L2, there was a re-arrangement of the L2 protein within the HPV16 PsV capsid. The atomic resolution structure of the HPV16 L1 only asymmetric unit (PDB ID: 5kep) (Figure 1.1[C]) was used to generate an L1-only density map in Chimera. To achieve this, icosahedral symmetry was applied using the 'sym' command in Chimera, generating sixty copies of the L1-only asymmetric unit, forming an entire HPV16 capsid. Following this, the command 'molmap' was used to generate a 3D density map of only the L1 protein at 5 Å and 7.3 Å respectively. Then the two HPV16 PsV density maps generated in this project were imported, furin cleaved map was fitted into the 5 Å L1 only density map, while the furin uncleaved map was fitted into the 7.3 Å map. The command 'vop subtract' was used to subtract either the furin uncleaved or furin cleaved HPV16 PsV map from the 7.3 Å or 5 Å L1-only map respectively. The remaining density should only correspond to density associated with L2 (50, 55).

3 Results

Investigation into the entry mechanisms used by HPV for infection of host cells may aid in the development of new approaches to prevent infection and decrease the risk of HPV linked cancers. After HPV attaches to cell surface HSPGs, a host protease (furin) cleaves the minor HPV capsid protein, L2. Furin cleavage of L2 is a necessary step in HPV infection, without which the virus is far less infectious (sections 1.5.2 and 1.5.3) (106, 110, 116, 118). Therefore, detailed investigation of the effects of furin cleavage on the HPV capsid may present new avenues for therapeutic development. Furthermore, it has been proposed in the literature that changes to the HPV capsid structure after furin cleavage prime the virus for entry into host cells, although this has never been directly visualised. This may be achieved by reducing the virus's affinity to HSPGs, or a structurally similar substitute molecule, heparin, allowing virus transfer to the entry receptor(s) (section 1.5.3). This work made use of HPV16 PsVs as they are functionally and morphologically similar to native HPV16 and have been successfully studied using structural biology (41, 49, 50, 102, 105). HPV16 PsVs were used to directly visualise structural changes to the HPV16 PsV capsid upon furin cleavage and the associated consequences for HPV binding to heparin.

3.1 HPV16 PsV production and evaluation

Before use in *in vitro* and resulting computational analysis, the purity and infectivity of the HPV16 PsVs were evaluated. Denaturing SDS-PAGE followed by silver staining revealed uncontaminated HPV16 PsV preparations and a prominent L1 protein band at 55 kDa associated with a co-migrating L2 protein band at 75 kDa in pure preparations of the HPV16 PsVs (Figure 3.1[A]). HPV16 PsV infection of HaCaT cells, after virus pre-incubation with either the HPV16 specific antibody (H16.V5) or the HPV18 specific antibody (H18.J4; negative control) (section 2.5.3) revealed complete abolishment of infection using HPV16 PsV bound with H16.V5 antibody, but no decrease in infection using HPV16 PsVs pre-incubated with H18.J4 (Figure 3.1[B]). Finally, negative staining and TEM visualisation (section 2.7) of purified HPV16 PsV capsids revealed homogenous, spherical ~55 nm capsids (Figure 3.1[C]).

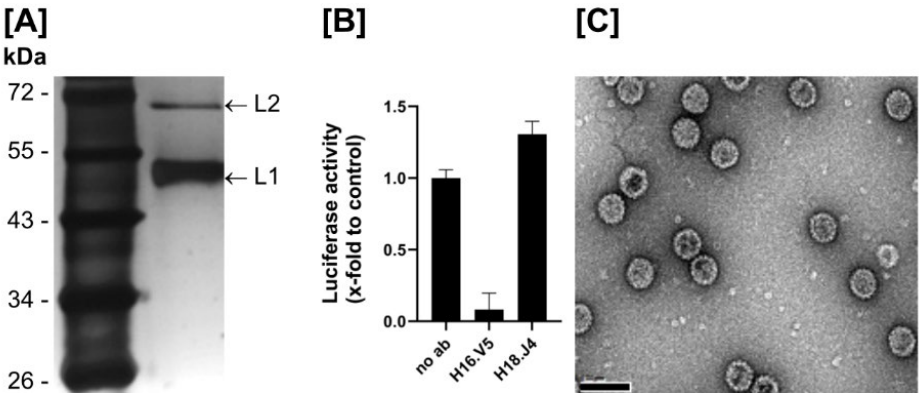


Figure 3.1. HPV16 PsVs are uncontaminated and display spherical capsids. [A] Silver stained SDS-PAGE gel showing purity of HPV16 PsVs preparation. A thick protein band corresponding to the L1 protein (55 kDa) and a thin L2 band (75 kDa) of HPV16 were seen in the preparations used in this project. **[B]** HPV16 PsVs (encapsidating a Gaussia luciferase reporter gene) were pre-incubated in the presence or absence of an HPV16 (H16.V5) or an HPV18 (H18.J4) neutralising antibody and added to HaCaT cells. Luciferase values are presented as x-fold change relative to HPV16 PsV infection of HaCaT cell in the absence of any antibodies, which was set to 1. Bars depict mean values, while error bars indicate standard error of the mean. **[C]** Negative stain and transmission electron microscopy of purified HPV16 PsVs, demonstrated spherical and homogenous ~55 nm HPV16 capsids. Scale bar 100 nm.

3.2 The structure of HPV16 PsVs

After determining that the HPV PsV preparations were uncontaminated and infectious, the particles were vitrified onto holey thin-carbon coated copper grids and imaged by cryo-EM (section 2.9 and Table 3.1). Cryo-EM was done to confirm whether the HPV16 capsid can be successfully imaged by cryo-electron microscopy and lead to a 3D reconstruction, before further investigation. The HPV16 PsVs were ~55 nm size, homogenous and well distributed within the holes of the cryo-EM grid (Figure 3.2[A]). The collected dataset resulted in a 3D reconstruction at 34 Å resolution, which displayed the ~55 nm spherical diameter and “knobby” spherical capsid exterior, made up of 72 capsomeres with icosahedral symmetry (Figure 3.2[B] and Table 3.1). These capsomeres appeared pentameric, as would be expected if each capsomere was composed of five L1 proteins. This is confirmed by ‘rigid-body’ docking of a PDB made up of the L1 capsid protein (PDB ID: 5kep) structure into our cryo-EM map (section 2.15) (50). The Fourier Shell Correlation (FSC) curve demonstrated the achieved resolution at 34 Å, where the curve crosses the FSC threshold of 0.143, the ‘gold- standard’ for determining density map resolution (Figure 3.2[C]) (178, 179). These results demonstrate the HPV16 PsVs produced are of good quality and could be further studied by high-resolution cryo-EM.

Table 3.1. Cryo-EM data collection parameters and refinement statistics for the low-resolution furin uncleaved HPV16 PsV dataset.

Data collection parameters	
Microscope	F20
Detector	US 4000 CCD camera (Gatan)
Voltage (kV)	200
Pixel size (Å)	2.11
Nominal magnification	53 000
Defocus (μm)	-0.3
Number of images collected	77
Microscope parameters	
Cs (mm)	1.22
Map reconstruction	
Initial particle number	755
Final particle number	681
Particle box size (pixel)	600
Resolution (gold standard FSC = 0.143) (Å)	34
Map sharpening β factor ($Å^2$)	0
Solvent mask (Å)	600

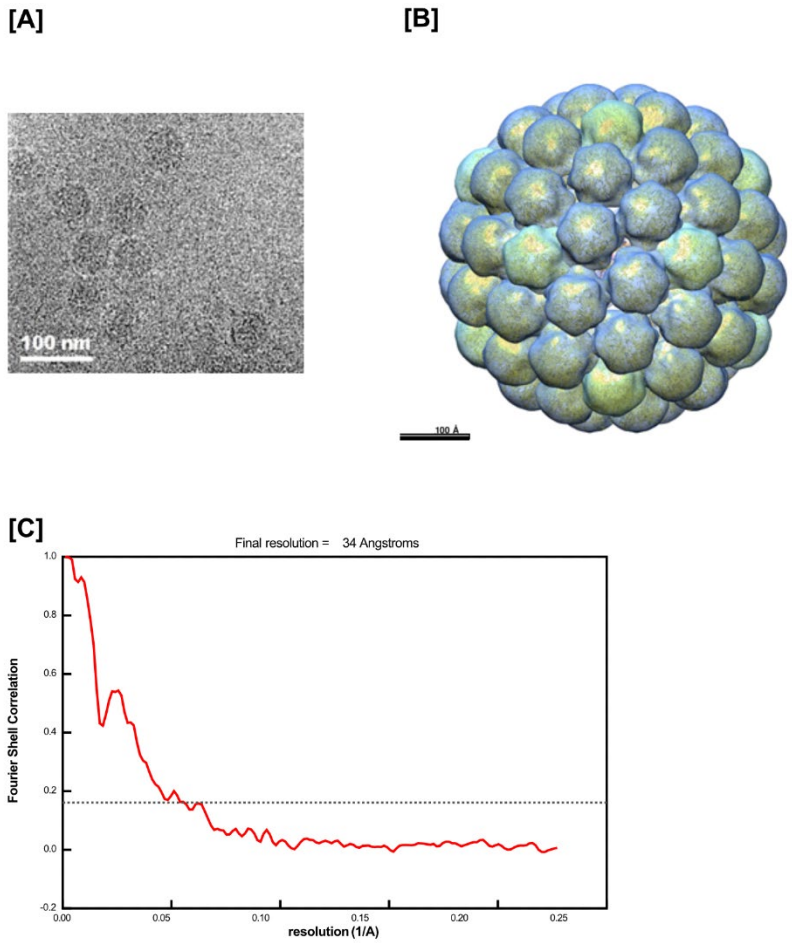


Figure 3.2. Cryo-electron microscopy (cryo-EM) of HPV16 PsVs confirm structure of the HPV16 pseudovirion capsid. [A] Micrograph of HPV16 PsV imaged by cryo-EM bound to holey thin carbon. Scale bar 100 nm. **[B]** Three-dimensional reconstruction of HPV16 PsVs demonstrating the pentameric capsomeres, icosahedral symmetry and spherical HPV16 PsV capsid. The L1 protein model (PDB ID: 5kep) was docked into the 3D density map (50). **[C]** Fourier Shell Correlation curve demonstrating achieved resolution in HPV16 PsV density map. The resolution the map was assessed where the Fourier Shell Correlation curve crossed a correlation value of 0.143, according to the 'gold-standard' (dotted line).

3.3 Production and purification of furin cleaved and uncleaved HPV16 PsVs for high resolution cryo-EM

HPV16 PsVs were cleaved by the addition of exogenous furin (section 2.4) to enable direct visualisation of the resulting physical changes to the HPV capsid after furin cleavage. Additionally, furin uncleaved HPV16 PsVs were used as a negative control for the effects of furin on the HPV16 capsid. Successful capsid protein expression and furin cleavage of L2 was demonstrated by SDS-PAGE and silver staining, indicating a double L2 band at ~75 kDa within the furin processed PsVs, compared to the furin unprocessed HPV16 PsVs (Figure 3.3[A]). Incomplete furin cleavage appears to be the norm (105, 126). Subsequently, the furin cleaved HPV16 PsVs were tested for infectivity in HeLa cells. HeLa cells infected with either furin uncleaved and cleaved pseudovirions exhibited higher luciferase readings than uninfected cells, indicating the luciferase reading was not due to background signal (Figure 3.3[B]). Moreover, virions cleaved by exogenous furin were much more infectious than uncleaved virions (Figure 3.3[B]). High-resolution cryo-EM micrographs of furin cleaved and uncleaved HPV16 capsids displayed spherical virions with capsomere protrusions (Figure 3.3[C] and [D]). These micrographs also demonstrated that both furin cleaved and uncleaved HPV16 PsVs were structurally homogenous and maintained the expected spherical 55 nm capsid structure. These results complete objective 1.1.

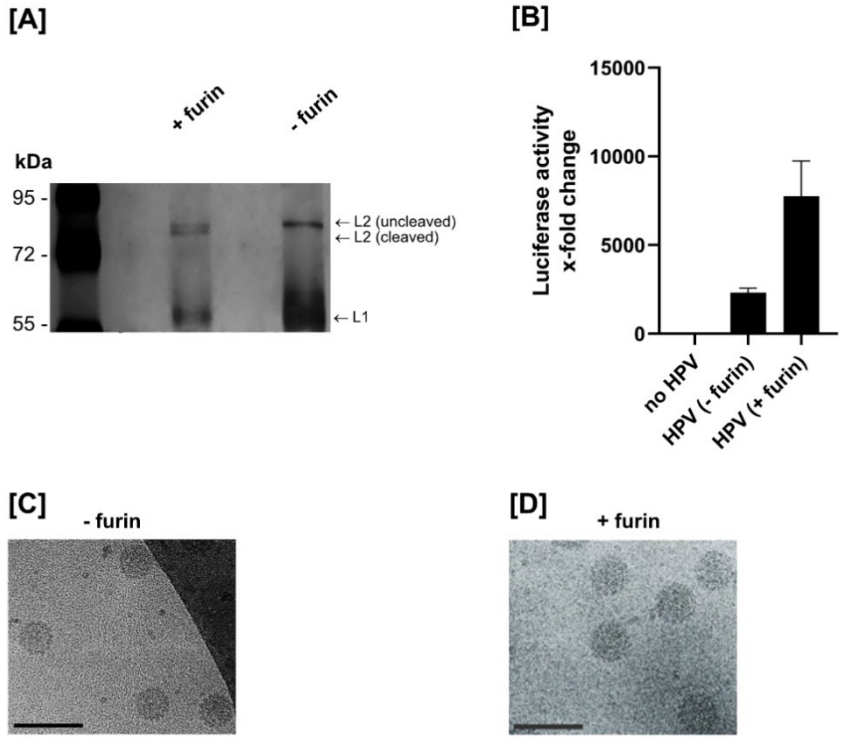


Figure 3.3. Successful generation and evaluation of furin cleaved HPV16 PsVs. [A] SDS-PAGE demonstrating successful furin cleavage of L2 at ~ 75 kDa, illustrated by the double L2 band '(+ furin)' lane, compared to the single L2 band in the '(-furin)' lane. **[B]** HeLa cells were infected with either furin cleaved or uncleaved HPV16 PsVs and Gaussia luciferase luminescence was measured 48 hours later as a read-out for infection. Luciferase values are presented as x-fold change relative to uninfected HeLa cells, which was set to 1. Bars depict mean values, while error bars indicate standard error of the mean. Cryo-EM images of **[C]** furin uncleaved HPV16 PsVs and **[D]** furin cleaved HPV16 pseudovirions illustrate homogenous, spherical ~55nm capsids with no distinguishable structural differences between furin uncleaved and furin cleaved HPV16 PsVs. Scale bar 100 nm.

3.4 High-resolution cryo-EM density maps

After sample vitrification, high-resolution imaging (section 2.11), and single particle analysis (sections 2.12 and 2.13), we produced two 3D reconstructions of the HPV16 PsV capsid; a furin uncleaved HPV16 PsV reconstruction (Figure 3.4 [A, C] and a furin cleaved HPV16 PsV (Figure 3.4 [B, D]) reconstruction at 7.29 Å and 5 Å, respectively (Table 3.5 and Table 3.6). The single-particle analysis and intermediate structures are presented in Figure 2.1 and Figure 2.2. The resolution values were calculated according to the 'gold-standard' Fourier Shell Correlation (FSC) (Figure 3.4 [C] and [D]) (178, 179). The density maps display a continuous 'shell' of protein density made up of pentameric hexavalent or pentavalent capsomeres, on a T=7 spherical capsid with a diameter of ~ 55 nm (Figure 3.4 [A] and [B]). Both capsids contain 72 capsomeres arranged with icosahedral two-fold, three-fold and five-fold symmetry. The capsids were composed of sixty asymmetric units and no differences in capsid diameter were observed, when comparing the furin cleaved and uncleaved HPV16 capsid (Figure 3.4 [A] and [B]). At the achieved resolution of 5 Å and 7.29 Å, some secondary structural features could be resolved in both maps and both reconstructions were at sufficient resolution to determine the position, orientation, and conformation of each L1 chain (PDB ID: 5kep) within the capsid structure (Figure 3.5). Alpha helices and beta sheets within each L1 chain (PDB ID: 5kep) corresponded well to the density in both furin cleaved and uncleaved reconstructions (Figure 3.5).

Table 3.2 Cryo-EM data collection parameters, refinement and validation statistics for the furin uncleaved HPV16 PsV dataset.

Data collection parameters	
Microscope	Titan Krios
Detector	K3
Voltage (kV)	300
Pixel size (Å)	0.53
Nominal magnification	81 000
Defocus range (μm)	-0.75, -1.0, -1.25,-1.5,-1.75,-2.0
Total dose ($e^-/Å^2$)	45
Exposure time (s)	4.01
Dose per fraction ($e^-/Å^2$)	1
Number of images collected	9 464
Number of frames per image	45
Microscope parameters	
C2 aperture (μm)	70
Objective aperture (μm)	100
Cs (mm)	2.7
Map reconstruction	
Initial particle number	20 595
Final particle number	3 913
Particle box size (voxel)	750
Resolution (gold standard FSC = 0.143) (Å)	7.29
Map sharpening β factor ($Å^2$)	-20
Solvent mask (Å)	700 or 800

Table 3.3 Cryo-EM data collection parameters, refinement and validation statistics for the furin cleaved HPV16 PsV dataset.

Data collection parameters	
Microscope	Titan Krios
Detector	K2
Voltage (kV)	300
Pixel size (Å)	1.048
Nominal magnification	130 000
Defocus range (μm)	-0.75, -1.0, -1.25,-1.5,-1.75,-2.0
Total dose ($e^-/Å^2$)	45
Exposure time (s)	7.5
Dose per fraction ($e^-/Å^2$)	1
Number of images collected	4 162
Number of frames per image	45
Microscope parameters	
C2 aperture (μm)	70
Objective aperture (μm)	100
Cs (mm)	2.7
Map reconstruction	
Initial particle number	20 549
Final particle number	3 913
Particle box size (voxel)	750
Resolution (gold standard FSC = 0.143) (Å)	5
Map sharpening β factor ($Å^2$)	-95
Solvent mask (Å)	700

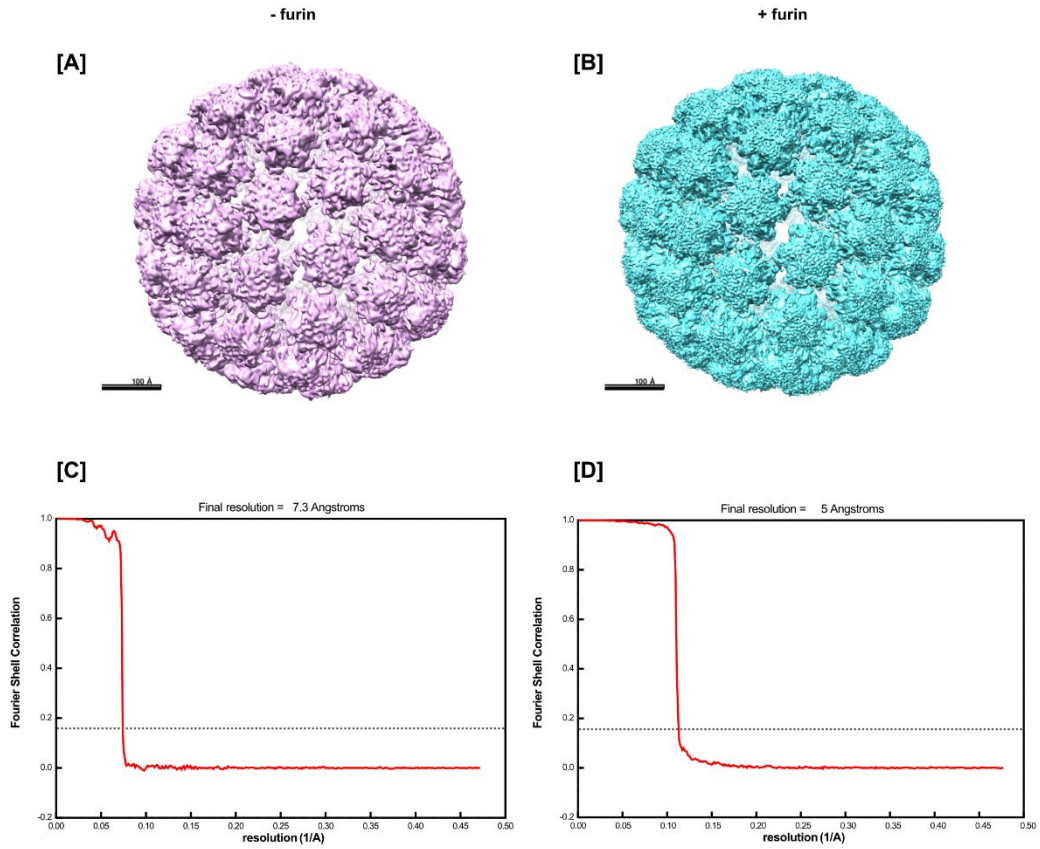


Figure 3.4. Structure of HPV16 PsVs imaged by cryo-electron microscopy (cryo-EM) achieved high- resolution. Surface rendered three-dimensional density maps of **[A]** furin uncleaved (- furin) and **[B]** furin cleaved (+ furin) HPV16 PsVs confirm spherical capsid structure and pentameric capsomers arranged according to icosahedral symmetry. No difference in capsid diameter after furin cleavage was observed. Scale bar 100 Å. Fourier Shell Correlation curve demonstrating achieved resolution in the **[C]** furin uncleaved and **[D]** furin cleaved HPV16 PsV density maps. The resolution of each map was assessed where the Fourier Shell curve crossed a correlation value of 0.143, according to the 'gold-standard' (dotted line).

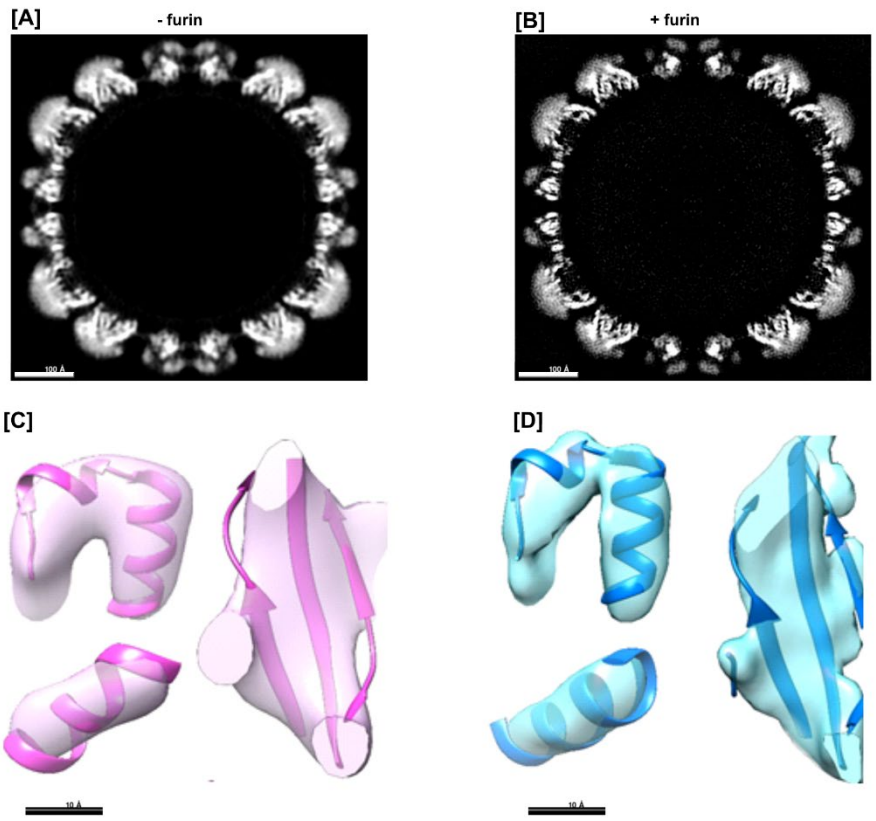


Figure 3.5. Secondary structures within the L1 capsid protein were identified in both density maps. Central sections through the cryo-EM density map of **[A]** furin uncleaved and **[B]** furin cleaved HPV16 PsVs, illustrate the quality of the reconstructions. The outer surface is less defined relative to the internal surface. Scale bar 100 Å. The atomic structure (PDB ID 5kep) of the L1 capsid protein was docked into the **[C]** furin uncleaved and **[D]** furin cleaved HPV16 PsV density maps, demonstrating that the resolution obtained was sufficient for determining the orientation and conformation of each L1 chain making up the capsid. Representative regions of secondary structures are shown to illustrate goodness of fit. Scale bar 10 Å.

3.5 Furin cleavage of L2 triggers a conformation change in the L1 chains within the HPV16 capsid

Cleavage of L2 by furin is suggested in the literature to enable conformation changes in the HPV16 capsid necessary for later infection steps (section 1.5.2), but this has never been directly visualised. There is an unknown number of L2 molecules within the HPV16 capsid and their location within the capsid is inconclusive (55, 130). Consequently, we could only assess structural changes to the L1 capsid protein. There were no obvious differences in capsomere structure, arrangement or capsid diameter identified between the furin uncleaved and furin cleaved density maps (Figure 3.4). To assess changes in the arrangement of the L1 proteins after furin cleavage, we rigid-body docked the individual L1 protein chains (PDB ID: 5kep) into either the furin uncleaved or furin cleaved HPV16 PsV density maps generated here, without alignment of the L1 chain A (section 2.16). After which, each L1 chain within both pentavalent and hexavalent capsomeres was in a different position after furin cleavage (Figure 3.6). The change in position of the L1 chains was independent of other L1 chains, and the whole capsomere does not rotate or translate in a particular direction (Figure 3.6). Chain A (forming part of the pentavalent capsomere) was then aligned to assess the conformation change relative to a fixed point. Each capsomere has an altered position relative to adjacent capsomeres, and the central cavity within each capsomere expands after furin cleavage (Figure 3.7 and Figure 3.8). The canyon between a pentavalent and hexavalent capsomere is reported to be the site of heparin binding (50). Therefore, this region on the HPV16 capsid is of particular importance for assessing the functional implications of a furin mediated capsid conformation change for later steps in HPV16 infection. We then focussed our investigation on this region of the HPV16 capsid, using the asymmetric unit of the HPV16 capsid (PDB ID: 5kep; Figure 1.2), as this represents all the positions of the L1 chains on the HPV16 capsid. To assess the change in capsomere conformation after furin cleavage, the asymmetric unit in the furin uncleaved and cleaved conformation were superimposed, and chain A was aligned between the two conformations (Figure 3.7[A]). The RMSD was calculated for each chain (chains A to F) (Figure 3.7[A] and Table 3.4). Across all six L1 chains, the RMSD was 0.940 Å, with chain E displaying the greatest movement, at 1.677 Å and chain C displaying the least movement after furin cleavage, at 0.762 Å (Table 3.4). In the putative heparin binding site, the L1 chains shifted on average 0.4 Å after furin cleavage (Figure 3.9). These results complete objective 1.2 and 1.3.

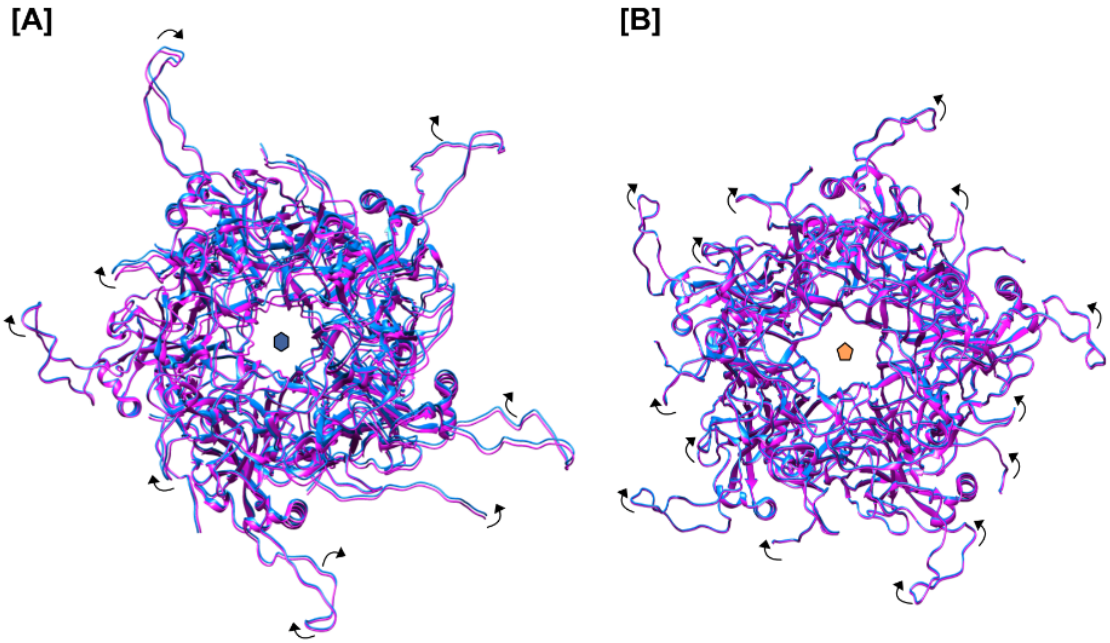
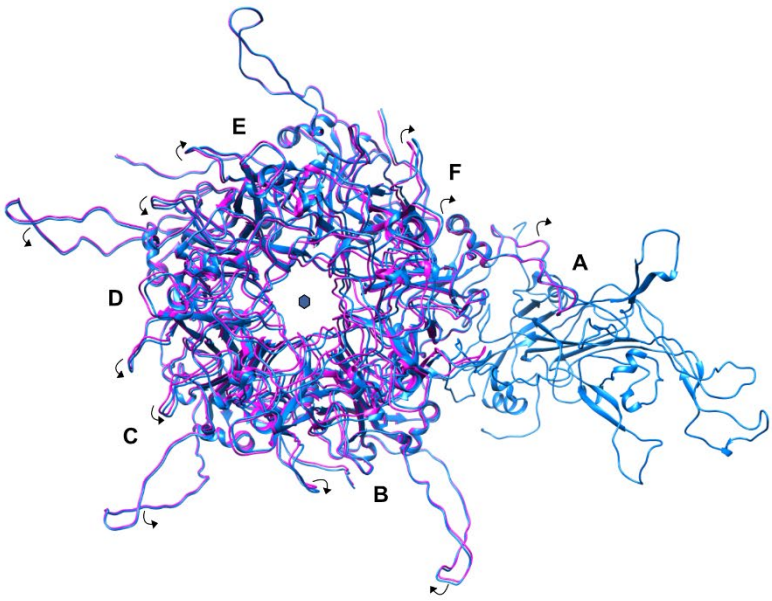


Figure 3.6. Atomic model reveals changes in the positioning of L1 chains within both pentavalent and hexavalent capsomeres after furin cleavage. Each L1 chain within a published L1-only HPV16 asymmetric unit (PDB ID 5kep) was docked individually into **[A]** hexavalent and **[B]** pentavalent capsomere position within the furin uncleaved (blue) and furin cleaved (magenta) density maps generated here. This was done without alignment of chain A (forming the pentavalent capsomere), demonstrating that both the pentavalent and hexavalent capsomere display altered conformations after furin cleavage. Black arrows demonstrate the movement of the L1 chains after furin cleavage. The blue hexagon indicates a hexavalent capsomers, while the orange pentagon indicates a pentavalent capsomere.

[A]



[B]

Figure 3.7. Atomic model reveals changes in the positioning of L1 chains after furin cleavage occurs in the HPV16 asymmetric unit. Each chain within a published L1-only HPV16 asymmetric unit (PDB ID 5kep) was docked into either the furin uncleaved (magenta) and furin cleaved (blue) conformation based on the density maps produced in this study. L1 chain A forming part of the pentavalent capsomere was aligned between the furin uncleaved and furin cleaved for quantitative assessment of the conformation changes of the other chains (B to F) after furin cleavage. **[A]** The L1-only asymmetric unit (PDB ID: 5kep) in either the furin uncleaved (magenta) or furin cleaved (blue) conformation. Each L1 protein chain within the asymmetric unit is labelled as A to F. Black arrows indicate the direction of movement of each L1 chain after furin cleavage. The blue hexagon indicates a hexavalent capsomere. **[B]** Video of the change in asymmetric unit (PDB ID: 5kep) conformation after furin cleavage. The L1-only asymmetric unit (PDB ID: 5kep) in either the furin uncleaved (magenta) or furin cleaved (blue) conformation. This video is also available at: https://vimeo.com/492103081.

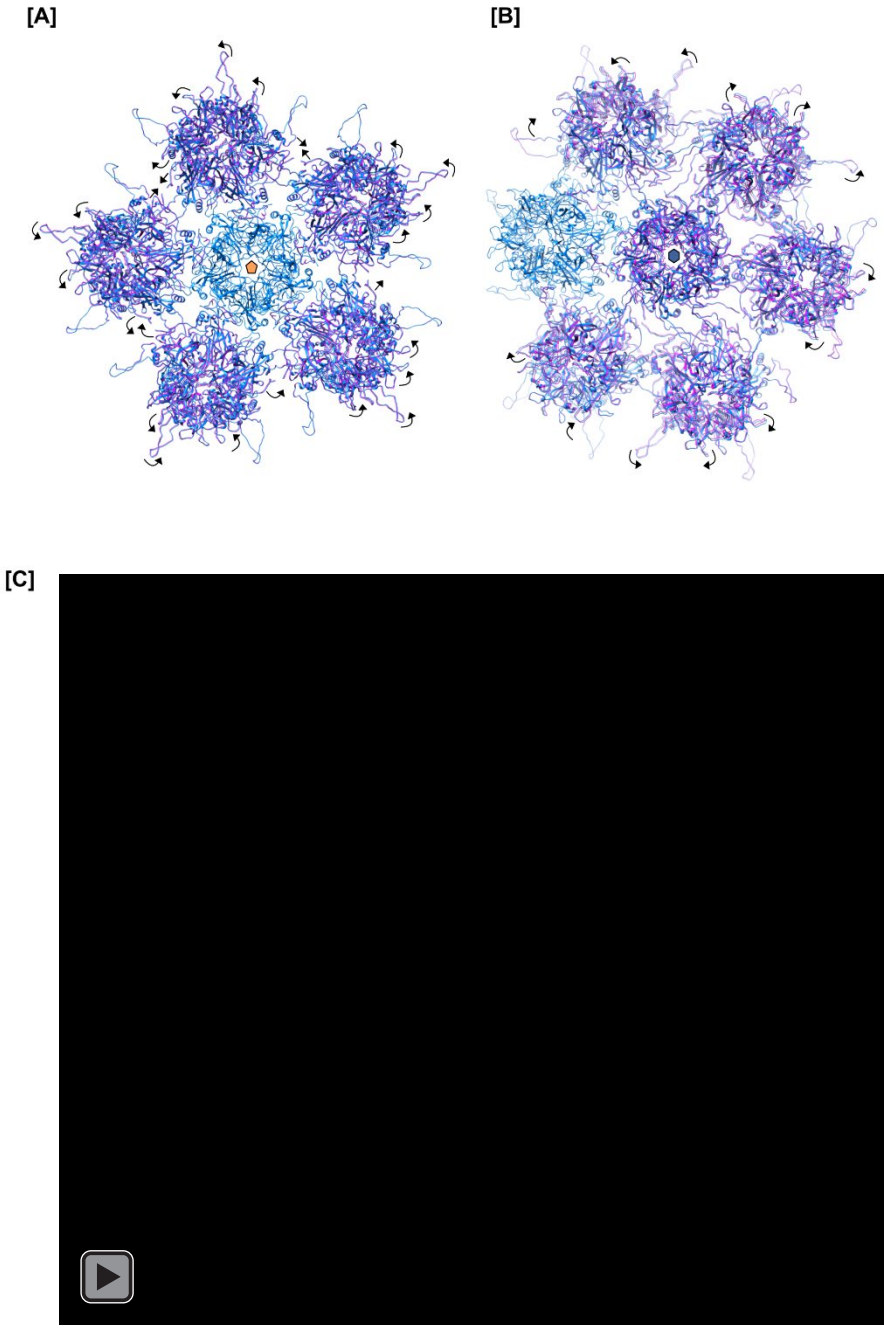


Figure 3.8. Atomic model reveals changes in the positioning of L1 chains after furin cleavage occurs across the HPV16 capsid. Each chain within a published L1 only HPV16 asymmetric unit (PDB ID: 5kep) was docked into the furin uncleaved (magenta) and furin cleaved (blue) conformation based on the density maps produced in this study. Chain A forming part of the pentavalent capsomere was aligned between the furin uncleaved and furin cleaved L1 capsomeres. **[A]** Pentavalent capsomere surrounded by five other capsomeres and **[B]** hexavalent capsomere surrounded by six other capsomeres, reveal the movement of L1 chains around these regions on the HPV16 capsid after furin cleavage. **[C]** Video of the capsid conformation change after furin cleavage. The furin uncleaved conformation is indicated in magenta, while the furin cleaved conformation is indicated in blue. This video is also available at https://vimeo.com/492078564,

Table 3.4. Quantification of HPV16 L1 chain movement after furin cleavage show each L1 chain display different degrees of movement.

Chain ID	RMSD (Å)
Chain A	0.000
Chain B	0.781
Chain C	0.762
Chain D	1.317
Chain E	1.677
Chain F	1.105
Average	0.940

3.6 HPV16 binds to heparin

After furin cleavage of L2, rearrangement of the L1 chains in the pentavalent and hexavalent capsomere was observed, including around the putative heparin binding site on the HPV16 capsid (Figure 3.9) (50). Therefore, we assessed whether the observed rearrangement of L1 chains after furin cleavage had any effect on HPV16 binding to heparin, using the molecular dynamics program ISOLDE (section 2.17) (180), the high-resolution HPV16 density maps produced here (Figure 3.4[A] and [B]). An anionic aspartic acid peptide was used as a substitute for heparin, as heparin is highly heterogenous, to broadly characterise the effects of furin cleavage on HPV16 binding to a negatively charged molecule. An overview of the results is presented in Figure 3.10. After running the simulation at 310 Kelvin, the anionic peptide formed hydrogen bonds to the side chains of the positively charged amino acids lysine (lys), arginine (arg) and histidine (his) on the L1 chains in the furin uncleaved and furin cleaved conformation (Figure 3.11, Table 3.5 and Table 3.6). Lys54, lys59, lys171, lys361, arg365, his431 on the L1 chains in furin uncleaved conformation all formed hydrogen bonds with the anionic peptide (Table 3.5), while on the L1 chains in the furin cleaved conformation lys59, lys430 and his431 formed hydrogen bonds with the anionic peptide (Table 3.6). Additionally, the peptide formed hydrogen bonds with polar amino acids on the L1 chain, including serine (ser), threonine (thr), asparagine (asn) and glutamine (gln), and non-polar amino acids including cysteine (cys), proline (pro) and isoleucine (ile) (Figure 3.11, Table 3.5 and Table 3.6). Ile60, ser173, cys175, thr176, asn177,cys365, cys428, gln429 and thr432 on the L1 chains in the furin uncleaved conformation formed hydrogen bonds with the anionic peptide, while thr176, thr432, asn177, gln429, cys428, cys175 on the L1 chains in the furin cleaved conformation formed hydrogen bonds to the anionic peptide (Table 3.5 and Table 3.6).

Notably, the L1 amino acids contributing to the aspartic acid binding originated from each of the three different L1 chains (chains A, B and F) around the canyon between the pentavalent and hexavalent capsomere. The majority of L1 amino acids binding the aspartic acid peptide were located on chain A, forming part of the pentavalent capsomere (Figure 3.11, Table 3.5 and Table 3.6). Fewer hydrogen bonds were observed between the L1 amino acids and the

anionic peptide after furin cleavage (compare Figure 3.11[A, B] to [C, D]; Table 3.5 to Table 3.6). These results complete objective 2.1.

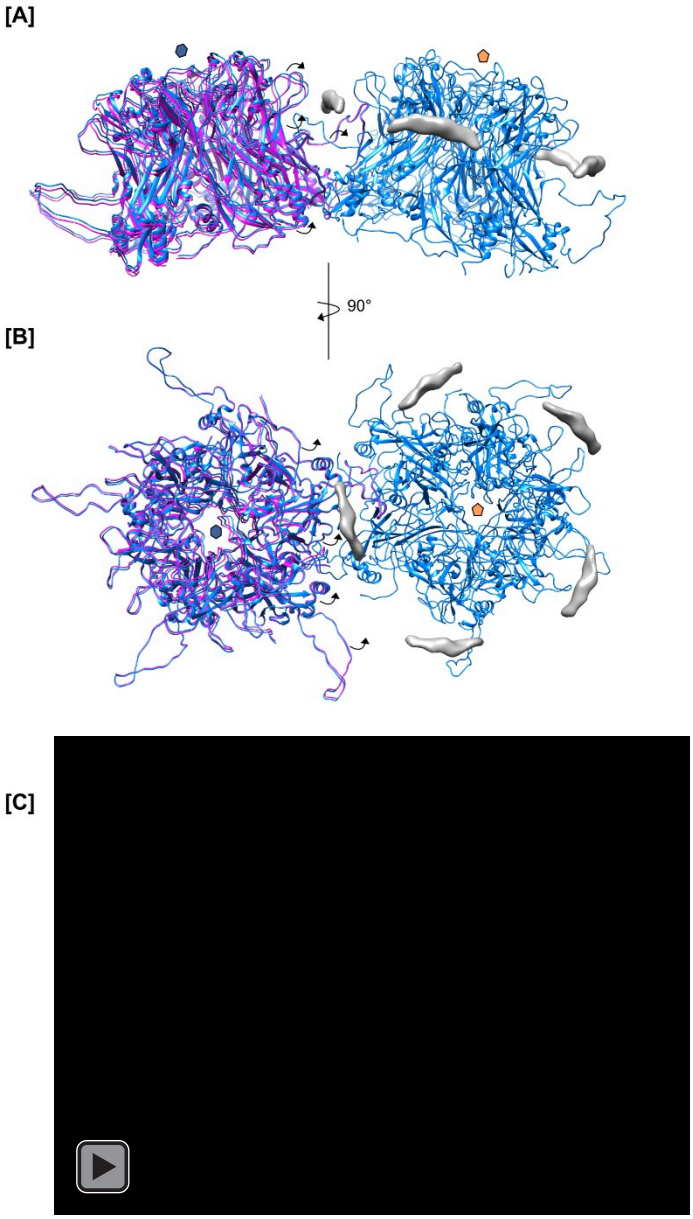


Figure 3.9. Furin cleavage of L2 causes a change in capsomere conformation around the putative heparin binding site. Cryo-EM suggests that heparin (surface rendered in grey) binds one of the (identical) L1 chains (chain A) in the pentavalent capsomere (indicated by an orange pentagon) and chains B and F within the hexavalent capsomere (indicated by a blue hexagon) (50). **[A,B]** The observed change in capsomere conformation, indicated by black arrows, between furin uncleaved (magenta) and furin cleaved (blue) also occurs around the heparin binding site. [C] Video of the conformation change after furin cleavage, around the heparin binding site. The furin uncleaved conformation is indicated in magenta, while the furin cleaved conformation is indicated in blue. This video is also available at https://vimeo.com/493342356.

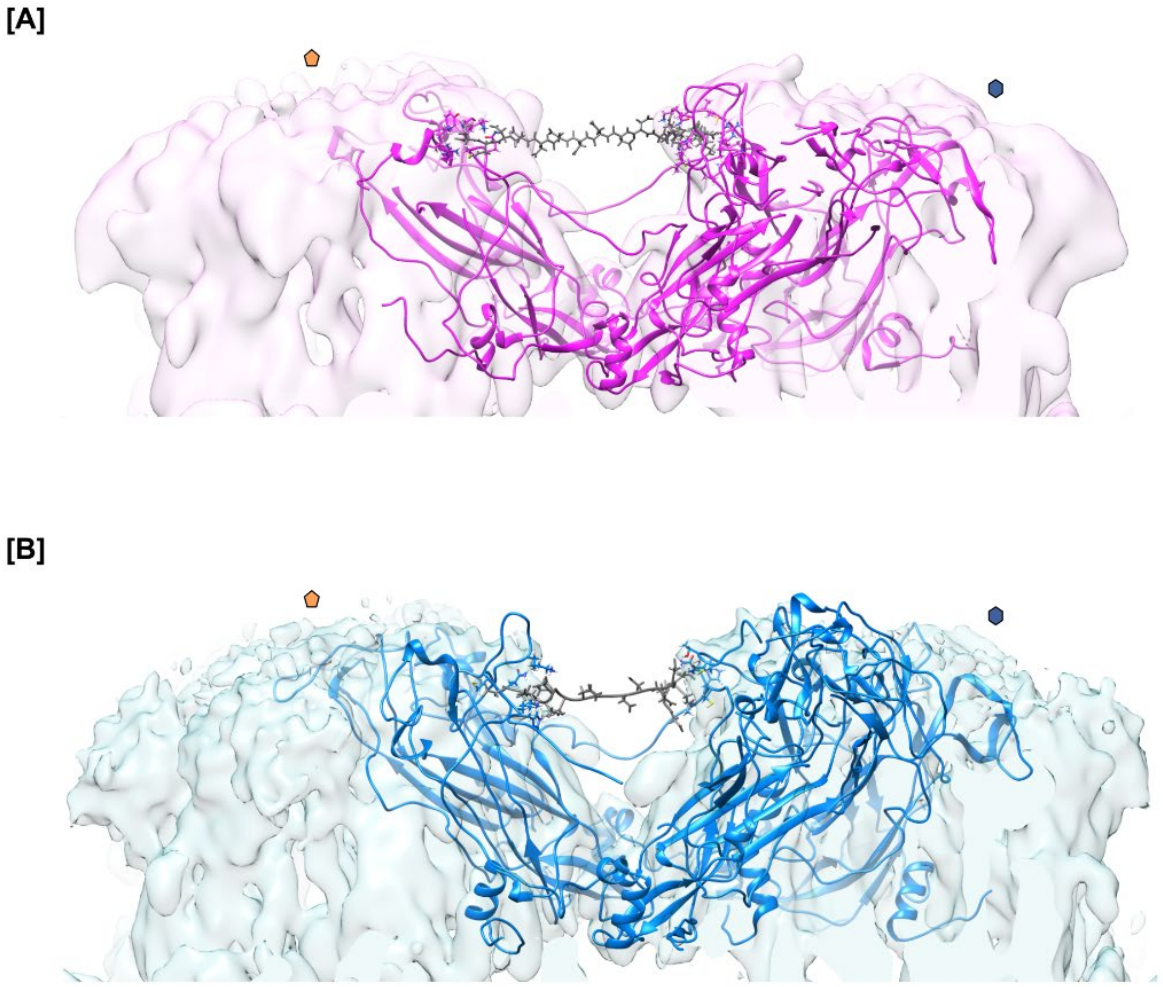


Figure 3.10. HPV16 L1 chains form hydrogen bonds with a highly negatively charged peptide.

Three L1 chains (chains A, B and F) of the HPV16 asymmetric unit (PDB ID: 5kep) fitted into **[A]** the furin uncleaved (magenta) and **[B]** the furin cleaved (blue) HPV16 pseudovirion density maps produced here. An anionic peptide (grey) was used as a heparin substitute and docked into the putative heparin binding site (50). Orange pentagon indicates a pentavalent capsomere, while the blue hexagon indicates a hexavalent capsomere.

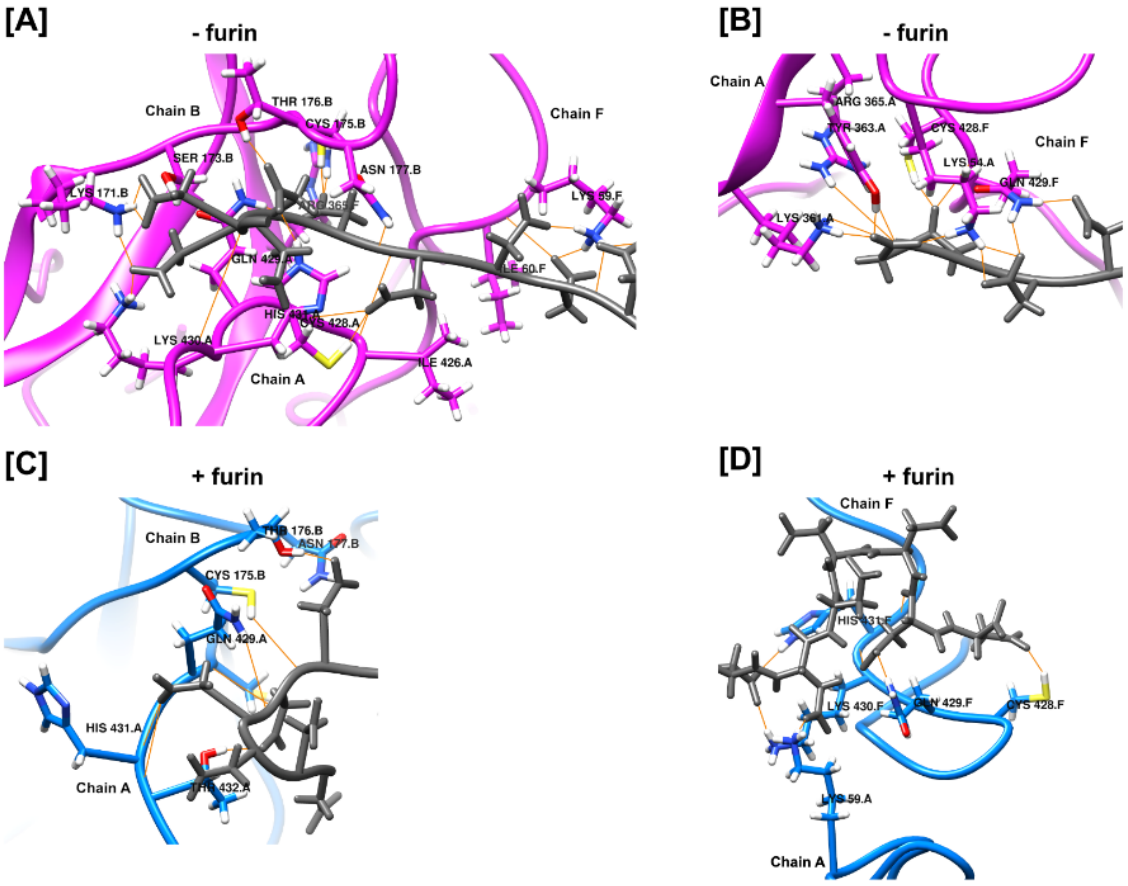


Figure 3.11. HPV16 L1 amino acids before and after furin cleavage forming hydrogen bonds with an anionic peptide docked into the putative heparin binding site. The three L1 chains (PDB ID:5kep) for the furin uncleaved HPV16 PsVs are indicated in magenta in **[A]** and **[B]**, while the furin cleaved HPV16 PsV L1 chains (PDB ID:5kep) are indicated in blue in **[C]** and **[D]**. The side chain of each interacting amino acid is shown and is labelled with the three letter residue name and L1 chain location (chain A, B or F). The aspartic acid peptide is coloured in grey while the hydrogen bonds are shown as orange lines between atoms. Part of the L1 chains A, B and F are labelled.

Table 3.5. HPV16 L1 amino acids interactions before furin cleavage of L2. Shown are the amino acids on HPV16 L1 proteins in the furin uncleaved capsid conformation, which interact with an aspartic acid peptide docked into the putative heparin binding site.

Chain A	**Chain B**	**Chain F**
lys54	lys171	lys59
lys361	ser173	ile60
tyr363	cys175	arg365
arg365	thr176	cys428
ile426	asn177	gln429
cys428	cys365	
gln429		
his431		
thr432		

Table 3.6. HPV16 L1 amino acids interactions after furin cleavage of L2. Shown are the amino acids on L1 proteins in the furin cleaved HPV16 conformation, which interact with an aspartic acid peptide docked into the putative heparin binding site.

Chain A	Chain B	Chain F
lys59	cys175	gln429
his431	thr176	cys428
cys428	asn177	lys430
gln429		his431
thr432		

3.7 Attempt at locating L2 within the furin uncleaved and furin cleaved density maps

Unfortunately, even at the resolution achieved, we were unable to resolve non L1 protein density (section 2.18) which may potentially correspond to the L2 protein, within either our furin uncleaved or furin cleaved density maps (data not shown). Given that no study to date has definitively demonstrated the location of the entire L2 protein within the HPV capsid, this result was not unexpected.

4 Discussion

Persistent infection with high-risk HPV types is associated with several cancers, especially cervical cancer, which is prevalent in areas of the world with high HIV/AIDS prevalence like South Africa (5, 8, 17, 20, 21). Prophylactic vaccines for HPV are available and safe but have some shortcomings including limited HPV type cross-protection, unknown long-term efficacy, cost, and accessibility (32, 33, 36, 44, 45). Despite considerable efforts by HPV vaccination programs both in South Africa and globally, cervical cancer remains a leading cause of cancer deaths in women worldwide (20). This highlights the need for the development of additional therapies to decrease the burden of HPV infection and HPV-linked cancers worldwide.

A potential therapeutic target is preventing HPV entry mechanisms used by the virus to infect cells. To initiate infection, HPV16 is primed for virus entry by several host proteins, including initially binding to HSPGs on the cell surface followed by enzymatic processing, most notably by furin. Without furin cleavage of the L2 capsid protein, HPV has far reduced infectivity in HeLa cells (Figure 3.3[B]) and is non-infectious in furin deficient cell lines. This indicates furin processing is an essential step in HPV infection (87, 99, 106, 109, 110, 112, 126).It is proposed, using antibody neutralisation assays, the furin cleavage of L2 facilitates a change in capsid structure, although this has never been directly visualised. Therefore, the present study focussed on direct visualisation of the effects of furin on the HPV16 PsV capsid. Clarifying the impact of this protease on the HPV16 capsid using structural biology may aid in the development of novel therapeutics against this prevalent and oncogenic virus. To achieve this, biochemical characterisation, cryo-electron microscopy, single particle analysis and molecular dynamics were applied in this study to describe the physical effects of furin cleavage on the HPV16 capsid. Here we describe the first furin cleaved HPV16 PsV structure, at 5 Å resolution.

4.1 HPV16 pseudovirions were suitable for high-resolution electron microscopy and single particle analysis

HPV16 PsV production and purification, followed by SDS-PAGE, silver staining, HeLa cell infection and HaCaT cell antibody neutralisation indicated HPV16 had been successfully produced and were suitable for downstream single-particle analysis (Figure 3.1[A] and [B]). The HPV16 PsVs which (per definition) encapsidated a pseudogenome (i.e. a reporter gene plasmid) exhibited T=7, spherical and ~55 nm capsid structures as seen for both pseudovirions and native HPV (4, 47, 49, 50, 67, 69, 126, 159, 160) displayed homogenous capsid morphology and were well distributed on the TEM grid (Figure 3.1[C], Figure 3.2[A]). We noted that the uranyl acetate negative stain penetrated some of the virions, causing them to appear ‘dented’ on micrographs (Figure 3.1[C]). These virions may be empty particles, containing no plasmid DNA and

therefore easily stained (49). However, the majority of HPV16 PsVs imaged were suitable for TEM imaging and single particle analysis. Apparent conformation changes to proteins due to adherence to plastic is a known issue in HPV research (128, 163). The risk of capsid conformation changes due to adherence to plastic was minimised by using siliconized Eppendorf tubes during pseudovirion production and the furin cleavage experiments. Taken together, the experimental conditions and virion quality was consistent with published results and demonstrate the purity, infectivity, and type-specificity of the HPV16 PsVs produced here (4, 47, 49, 50, 67, 69, 126, 159, 160).

The initial low-resolution HPV16 PsV reconstruction demonstrated the expected capsid structure and morphology, with 72 pentameric capsomeres, arranged as either hexavalent or pentavalent capsomeres and organisation around two-fold, three-fold, or five-fold icosahedral axes of symmetry (Figure 3.2[B]). Likewise, the high-resolution reconstructions of both furin uncleaved and cleaved HPV16 PsVs also display these characteristics (Figure 3.4[A] and [B]). Additionally, secondary structural features such as alpha helices and portions of the central β jelly-roll could be distinguished from the density in both high-resolution reconstructions (Figure 3.5[C, D]). L1 protein surface loops and C-terminus arms extending from each pentamer could not be resolved in either of the two high-resolution density maps produced here, as these are the most flexible regions within the HPV capsid (50, 55, 56, 183). Indeed, the less defined density on the capsomere surface of both the furin uncleaved and furin cleaved reconstructions likely reflect the inherent flexibility of the surface loops on individual capsomeres (Figure 3.4[A.B]).

After HPV16 binding to cell-associated HSPGs and capsid processing by host cyclophilins and kallikrein-8, furin cleaves the exposed N-terminus region on L2, reportedly leading to a change in the arrangement of the L1 and L2 capsid proteins (106, 110). Furin cleavage only removes about twelve amino acids off the L2 protein, which we visualised as a slightly lower band compared to uncleaved L2 on SDS-PAGE (Figure 3.3[A]). Incomplete furin cleavage is expected since complete furin cleavage of L2 has not been observed in the intact HPV capsid (99, 105, 106, 112, 126). The incomplete cleavage of L2 that we observed may be due to either all HPV16 capsids being partially cleaved by furin, or a subset of capsids being fully cleaved (112) Regardless, incompletely furin cleaved capsids still indicate that some of the L2 N-termini are exposed on the capsid surface and furin cleaved HPV16 PsVs were still highly infectious compared to uncleaved HPV16 PsVs (Figure 3.3[B]) (126). Becker *et al.* have demonstrated that only furin cleaved L2 proteins could be detected in the infected cell's nucleus, while furin uncleaved L2 proteins remain within the endosome. This indicates incomplete furin cleavage of L2 is sufficient for infection and that HPV16 PsVs cleaved by

exogenous furin have the same infectious pathway as HPV16 PsV infection of cells expressing furin and cell surface HSPGs (105). Therefore, even though we observed incompletely furin cleaved L2 within the HPV16 PsV capsid, this may represent a natural infection *in vivo*.

4.2 Conformational changes to the HPV16 capsid after furin cleavage

To investigate changes in the HPV capsid which may occur after furin cleavage of L2, we compared the two high-resolution reconstructions of furin uncleaved and cleaved virions produced here (Figure 3.4[A] and [B]), by rigid-body docking all six of the L1 chains in the atomic model of the HPV16 asymmetric unit (PDB ID: 5kep). This would allow us to interpret the differences in overall capsomere positioning between the two density maps to atomic resolution, providing us with better insight into the biological mechanism of HPV16 entry into host cells. The conformation changes to the L1 proteins visualized here after furin cleavage occurred across the HPV capsid, rather than stochastically or localised to a specific capsomere (Figure 3.8). Changes in quaternary structure across the entire viral capsid may explain the slow exposure of the L2 N-terminus and the observation that furin cleavage of L2 is the rate-limiting step in HPV entry (87, 105). The 0.940 Å change in the positioning of L1 chains after furin cleavage observed was not surprising as electron microscopy of furin uncleaved and furin cleaved HPV16 PsVs revealed no obvious changes in overall capsid structure when viewed by TEM (Figure 3.3[C, D] and Table 3.4), consistent with previous research (126, 127). Our results add direct evidence to the hypothesis that the furin cleaved HPV16 capsid is an intermediate structure between binding and entry (105).

Many viruses exploit furin for proteolytic processing during host cell entry, using furin mediated changes in capsid proteins to enhance infectivity, and HPV may be no exception (97, 106, 115, 120-123, 184). Indeed, there is indirect evidence that L1 protein chains within the HPV capsid undergo a conformation change after furin cleavage of L2, as we confirmed by cryo-EM (Figure 3.6, Figure 3.7, Figure 3.8, Figure 3.9 and Table 3.4) (87, 93, 105, 112, 127, 128, 132, 158). Most studies infer a conformation change in L1 after furin cleavage based on the RG-1 antibody detection of a previously unexposed region on L2 (amino acids 13-36). This region on L2 is only exposed after furin cleavage, as RG-1 only immunoprecipitated furin cleaved HPV16 PsVs. Likewise, HPV16 PsVs bound to HaCaT cells treated with a furin inhibitor could not bind RG-1, indicating furin is the vital host molecule necessary to expose the N-terminus region of L2. L1 protein(s) likely occlude the L2 N-terminus as two antibodies (H16.V5 and H16.E70) with epitopes on the HPV16 L1 protein prevent exposure of the RG-1 binding site, after the virus-antibody complex was added to cells, although the antibody did

not prevent HPV binding to the cell surface (87, 105, 112, 127, 128, 132, 158). Similarly, two other L1 specific antibodies, H16.56E and H33.J3, allow virion binding to HeLa cells and internalisation, but inhibit intracellular infection. These antibodies possibly achieve this by preventing furin-associated conformational changes to the HPV capsid necessary for later steps in infection (92). Unfortunately, these antibody binding studies do not distinguish between structural changes to the L1 proteins upon cell surface binding and the L1 conformation change after furin cleavage, making definitive conclusions difficult. Additional studies of the binding capacity of the antibodies mentioned above to furin cleaved HPV16 PsVs may help clarify this uncertainty.

Other cellular proteins such as kallikrein-8 are also responsible for changes in the capsid during HPV16 entry (section 1.5.2 and 1.5.3) (105, 132). It is therefore likely that the conformational changes visualised here after furin cleavage in combination with conformation changes induced by other host proteins prepare the virus capsid for internalisation and/or uncoating. Notably, furin cleaved HPV16 capsids are considered to be in the final conformation necessary for infection (56, 105).

The observed structural changes in the furin cleaved HPV16 PsV capsid obtained here occur both within capsomeres and between capsomeres (Figure 3.6, Figure 3.7, Figure 3.8, Figure 3.9 and Table 3.4). The exact function of the conformation change in the L1 proteins is presently unknown. However, the region of the L2 N-terminus which is now exposed is important for intracellular infection steps, and some research has found that this region on L2 is important for binding to the secondary cell receptor, although this is still inconclusive (110, 130, 131). Alternatively, the L1 conformation change after furin cleavage of L2 may begin the viral capsid uncoating process at the cell surface (103). An additional function of the L1 conformation change after furin cleavage of L2 may be that this processing affects the affinity of the virus capsid for HSPGs (or heparin), allowing the transfer of the HPV16 capsid to the elusive entry receptor or receptor complex (92, 93, 103). Indeed, HPV VLPs disassembled into individual capsomeres displayed no binding to cells, indicating that the entire capsid with intercapsomere contacts must be necessary for infection (89). Potentially the changes in capsomere structure and altered position of any capsomere relative to its' neighbours obtained here may affect the affinity of the capsid for the cell surface HSPGs, facilitating internalisation.

4.3 HPV16 binds to an anionic peptide within a putative heparin binding site

Heparin is a widely accepted substitute molecule for structural biology studies of the interaction between cell surface HSPGs and HPV (41, 50, 102, 103). Unfortunately, both heparin and HSPGs are flexible and heterogeneous molecules which do not have molecular dynamics force fields within programs such as ISOLDE. To circumvent this, we generated an anionic peptide, to simulate the negative surface charge and length of heparin, which was rigid body docked into a possible heparin-binding site in the canyon between hexavalent and pentavalent capsomeres (section 2.17) (50). In both the furin uncleaved and furin cleaved conformations, the amino acids on L1 interacting with the negatively charged peptide formed binding patches composed of more than a single L1 chain. These L1 chains originated from both hexavalent (chains B and F) and pentavalent (chain A) capsomeres (Figure 3.10, Figure 3.11, Table 3.5 and Table 3.6) similar to heparin binding results obtained by others (50, 102, 103).

The majority of L1 amino acids binding the anionic peptide in both furin cleaved and uncleaved structures were located on chain A, corroborating with heparin density located around the pentavalent capsomere (50). Furthermore, many of the of L1 amino acids forming binding patches to the anionic peptide were positively charged or polar amino acids (Figure 3.11, Table 3.5 and Table 3.6). In the L1 chains of the furin uncleaved conformation, we identified several amino acids forming hydrogen bonds to the anionic peptide, including lys59, thr176, and asn177 originating from the L1 chain within the hexavalent capsomere and cys428, gln249, his431 and thr432 from the L1 chain within the pentavalent capsomere (Figure 3.11 and Table 3.5). The specific amino acids listed here are identical to the HPV16 L1 amino acids interacting with heparin identified by Guan *et al.* (50). This confirms the suitability of the anionic peptide as a heparin substitute. Our results in combination with Guan *et al.* suggest these amino acids are required for viral capsid binding to HSPGs (50). Our molecular dynamics result also proposes that HPV16 binding to heparin (and potentially HSPGs) depends on the arrangement of these interacting L1 amino acids in relation to one another and overall charge distributions, as suggested by Knappe *et al*. (69), rather than the competing hypothesis that defined structural features on heparin and HSPGs mediate HPV binding (93).

Contrary to our results and results obtained by Guan *et al*. (50), x-ray crystallography studies suggest heparin binds to the surface and vertically to the side wall of the capsomere, rather than horizontally between capsomeres (50, 102). As we presumed heparin binds horizontally in the canyon between capsomeres, it is not surprising that our results did not correlate to those obtained by x-ray crystallography. The x-ray crystallography studies made use of HPV

pentamers, rather than the entire HPV capsid. Consequently, it is unknown whether the described binding sites occur on the hexavalent or pentavalent capsomeres, and intercapsomere contacts potentially relevant for heparin binding may be absent (102). Interestingly, although the anionic peptide was docked horizontally into the heparin density prior to molecular dynamics during our investigation, the peptide was free to move within the canyon between capsomeres during the simulation. But we did not observe the peptide binding vertically to the L1 chains in either the furin cleaved or uncleaved conformation (Figure 3.10). However, initially docking the peptide horizontally into the heparin density may have biased the results. A further simulation with the peptide docked vertically along the capsomere wall may resolve this query.

As atomic level crystallography studies (102, 103) indicate several heparin-binding sites on the capsomeres, the heparin density resolved by cryo-EM (50) may represent the average density of all possible binding sites of heparin to the capsids. Consequently, discrepancies between the amino acids involved in heparin-binding between the crystal structures and published cryo-EM structure may be explained by this. Moreover, it is proposed that HPV16 makes use of several binding sites on L1 during entry and infection, therefore a combination of the heparin binding amino acids identified from crystal studies and cryo-EM studies might be used for successful HPV infection (102, 103). Our results also support this conclusion, as the simulation revealed two anionic peptide binding sites (one on either side of the intercapsomere ‘canyon’) in both the furin uncleaved and furin cleaved L1 conformation structures (Figure 3.9 and Figure 3.10). Consequently, the structural changes to the L1 capsid proteins observed after furin cleavage may reduce HPV16 binding to the initial HSPG cell surface receptor, allowing transfer to secondary HSPG sites on the cell, possibly forming part of the entry receptor complex required for entry.

4.4 Furin cleavage may reduce HPV16 binding to heparan sulphate proteoglycans

To assess whether furin cleavage alters the affinity of the L1 amino acids for the anionic peptide (the HSPG substitute molecule), we measured hydrogen bond formation between the L1 amino acids and the anionic peptide before and after furin cleavage. Our molecular dynamics analysis indicated that after furin cleavage the amino acids originating from L1 chains in the furin cleaved conformation had fewer hydrogen bonds to the anionic peptide, when compared to the amino acids on the L1 chains within the uncleaved conformation (compare Figure 3.11[A, B] to [C, D]; and Table 3.5 to Table 3.6). This result proposes that HPV16 has decreased affinity for the primary cell surface HSPG receptor after furin cleavage. In line with this, HPV16 PsVs processed by furin have reduced affinity to heparin and were found to dissociate from HSPGs on COS-7 cells (92, 105). Additionally, furin cleaved HPV16 PsVs can infect HSPG deficient cell lines, illustrating HSPG independent infection after furin cleavage (99, 126).

The molecular dynamics results also suggest that after furin cleavage, HPV16 PsVs make use of different L1 amino acids to interact with the negatively charged peptide (compare Figure 3.11[A, B] to [C, D]; and Table 3.5 to Table 3.6), although these L1 amino acids are all located on the L1 chains surrounding the intercapsomere canyon (Figure 3.10). This result is similar to other research suggesting HPV16 may attach to heparan sulphate molecules using different binding sites on the capsid to achieve either initial attachment to cells or later steps in the entry process (92, 102, 103). Consequently, molecular dynamics results also propose HPV16 does not dissociate from the HSPG substitute, after furin cleavage, but rather uses different amino acids on the L1 chains for subsequent steps in infectious entry (Figure 3.10 and Figure 3.11), as observed by others (92, 102, 103). Furin cleaved HPV16 PsVs are independent of HSPGs used for initial cell binding, compared to furin uncleaved virions (99, 109, 110). However, two other highly sulphated polysaccharides, carrageenan and heparin, both inhibited HPV16 infection (127). This result may indicate a further role of sulphated polysaccharides in infection after furin cleavage, although the initial cell surface binding HSPGs become unnecessary. Moreover, HPV33 VLPs bind strongly to HeLa cells, with a dissociation constant of 84 pM. This indicates a strong association, possibly by HPV33 binding to cell surface HSPGs. As a result, complete dissociation of the viral capsid after furin cleavage is less likely (89).

Even though all capsomeres on the HPV capsid are composed of five L1 chains (41, 50, 102), there are differences in the interactions between the two sets of capsomeres. The interface between a hexavalent and pentavalent capsomere is on the icosahedral five-fold axis of symmetry, where a single L1 chain of the pentavalent capsomere only contacts a single L1

chain on the hexavalent capsomere (Figure 1.1[B]). In contrast, the interface between the two hexavalent capsomeres is asymmetric, although it appears to have six-fold symmetry ('pseudo six-fold symmetry') (Figure 1.1[B]). Three L1 chains contact within the hexavalent capsomere interact with four neighbouring capsomeres, while the other two L1 chains within the hexavalent capsomere contact the remaining two surrounding capsomeres (Figure 1.1 [B]) (48, 50). These differences have functional consequences for HPV infection, as Guan *et al.* only resolved heparin density in the canyons between the pentavalent and hexavalent capsomere, and no heparin density within the canyon between two hexavalent capsomeres (Figure 3.9) (50). The C-terminus of L1 proteins, within the canyon between capsomeres, contain a cluster of basic amino acids, which may be used by the virus during attachment to cell surface HSPGs (47, 52, 69, 94) (94). Speculatively, the L1 C-terminus originating from hexavalent capsomeres may be less exposed in the canyons between hexavalent capsomeres when compared to the C-terminal arms originating from the pentavalent capsomere, due to the pseudo six-fold symmetry between two hexavalent capsomeres. Alternatively, the addition of a sixth capsomere around the hexavalent capsomere may disrupt HSPG binding sites which are present between a pentavalent and hexavalent capsomere. Interestingly, the majority of the L2 protein could be located within the pentavalent capsomere, with some regions (eg. The N-terminus) being surface exposed around the pentavalent capsomere (52, 60). It is conceivable that the surface exposed regions of L2 may also contact the cell surface HSPGs during virus entry, which may explain why heparin (or HSPGs) only binds around the pentavalent capsomere. Indeed, an L2 peptide composed of the N-terminus amino acids was found to bind to the cell surface and internalise into the cell (58, 131). However, other studies indicate that L2 is unlikely to be involved in cell-surface binding (59, 89). More studies of the possible interaction between HPSGs and the L2 capsid protein will be needed to clarify this.

As the identity, location and amount of the secondary entry receptor or receptor complex on the keratinocyte cell surface is speculative (section 1.5.3), it is unknown whether furin cleavage of L2 allows the virus to associate with this secondary receptor or receptor complex. Our study did not address this question but rather focused on structural changes to the HPV capsid before engagement with this receptor (complex). In addition, our results have the important caveat that a negatively charged peptide may not represent the exact sequence or flexibility associated with HSPGs, although we were able to replicate the negative surface charge on heparin, which is proposed as the primary determinant of HPV binding (69).

4.5 The location of L2

Unfortunately, this investigation could not resolve any non-L1 protein density which could be attributed to L2, within either the high-resolution furin uncleaved or furin cleaved reconstructions (data not shown). Even though L2 was present in HPV16 PsV capsids and successfully cleaved by furin (Figure 3.3[B]). Consequently, we could not resolve the effects of furin cleavage on the arrangement of the L2 protein within the HPV16 capsid. The exact location and number of L2 proteins within the HPV16 capsid is still inconclusive. This may be for several reasons including heterogeneous amounts of L2 in the HPV capsids or an asymmetric distribution of the L2 protein within a single capsid (60, 130). These considerations make determining the characteristics of L2 within the HPV16 capsid using cryo-EM difficult, as single particle analysis generates 2D and 3D class averages and often imposes icosahedral symmetry, which averages out density across the whole density map (169, 172). Consequently, any asymmetric or heterogenous density within or between imaged particles that could be attributed to L2 will be weaker in the final icosahedral and averaged reconstruction. Asymmetric distribution of L2 within the virus capsid may also explain the resolution limit obtained in our density maps (Figure 3.4) and those obtained by others (50, 60, 127). Focussed refinement of different class averages or sub-particle refinement may aid in determining the characteristic of L2 within HPV16 and improve map resolution. Recently (paper in the process of peer-review), this approach was applied to HPV16 and achieved a 3.1 Å density map and the location of some L2 residues was determined. These L2 amino acids were present in the interior both of hexavalent and pentavalent capsomeres around the L1 protein, although the entire L2 protein was not resolved. The authors suggest L2 has a symmetry mismatch in the number of L2 proteins per L1 protein, and individual capsomeres could not incorporate five copies of L2 per capsomere, as is the case for L1, due to capsomere size limitations (60).

The L2 N-terminus is apparently located in the canyon between the pentavalent and hexavalent capsomeres, as an antibody directed against L1 amino acids within this region, prevented binding of L2 specific antibodies (185). However, other studies propose the L2 N-terminus is located near the L1 surface loops on the outer side of the capsomere as an antibody directed against the L1 surface loops prevented furin cleavage of L2, while a second antibody directed against regions of L1 within the intercapsomere canyons, did not prevent furin cleavage of L2 (112). It is possible that the L2 N-terminus on the capsomere exterior is flexible and may be located between these regions, accounting for the discrepancies between these two studies. Further investigation will be needed to resolve this.

5 Conclusion and proposed mechanism

Our results in combination with evidence from other studies propose a model for HPV entry (Figure 5.1). Importantly, other host molecules and enzymes also play a role in successful HPV entry (sections 1.5.2 and 1.5.3), but for simplicity, we discuss the host molecules investigated here.

Figure 5.1. Simplified proposed model of HPV16 attachment and entry into host cells.

The HPV16 PsV (magenta) attaches to heparin (HSPG substitute molecule; surface rendered in grey) around the pentavalent capsomere. Furin cleaves the L2 minor capsid protein (not shown). Subsequently, the L1 chains on the HPV16 capsid undergo a conformation change, forming the furin cleaved HPV16 PsV capsid (blue). The HPV16 PsV then transfers to the secondary entry receptor or receptor complex, which may include the initial HSPG cell binding molecule (represented here by the heparin density) and is internalised into the cell. A video of this process is also available at https://vimeo.com/493004952/

HPV16 initially binds to cell surface HSPGs using binding patches composed of positively charged and polar amino acids on the side wall of both the hexavalent and pentavalent capsomere (Figure 3.9, Figure 3.10, and Figure 3.11). This induces a conformation change in L1 and furin cleavage of the L2 N-terminus (not visualised here). Furin cleavage causes further changes in L1 protein chains within both pentavalent and hexavalent capsomeres (Figure 3.6) This capsid conformation change reduces the number of hydrogen bonds

between the L1 amino acids and the anionic peptide acting as a heparin substitute. Additionally, the furin mediated conformation change allows hydrogen bonds to form between different L1 amino acids and the anionic peptide (Figure 3.11 [C] and [D]). These interactions may form part of the secondary receptor complex which includes cell surface HSPGs and facilitates internalisation. The conformation changes after furin cleavage may expose previously hidden secondary cell receptor binding sites on either L1 or L2 capsid proteins. It appears likely that a region on the L1 capsid protein binds the secondary receptor (complex) as L1 only VLPs have indistinguishable binding and internalisation patterns *in vivo* and *in vitro* compared to L1 and L2 containing VLPs (87). Moreover, HPV may use furin cleavage of L2 to expose conserved amino acids on L2 which are necessary for intracellular infection, rather than internalisation, only after the virus has successfully bound to host cells. This likely evolved as an immune evasion strategy to prevent the immune system from detecting the conserved L2 epitope before engagement with the host cell, where it may be too late for the elimination of the virus (128).

The results presented in this dissertation directly illustrate for the first time, that furin cleavage of L2 causes quaternary changes in the HPV16 capsid. This will help in the development of novel strategies to target HPV16 entry and infection, with the long-term aim to decrease the incidence of HPV associated cancers. Future work could be to apply a focussed refinement approach (60, 186) to both the furin uncleaved and furin cleaved HPV16 PsV reconstructions produced here, to possibly determine the effect of furin cleavage on the arrangement of L2 density within the capsid. Additionally, molecular dynamics using a heparin molecule docked into the putative heparin-binding site would also add further insight into the effects of furin cleavage on HPV's affinity for HSPGs.

6 References

The referencing method used in this dissertation was Vancouver.

1. Bravo IG, de Sanjose S, Gottschling M. The clinical importance of understanding the evolution of papillomaviruses. Trends Microbiol. 2010;18(10):432-8.

2. Chow LT, Broker TR, Steinberg BM. The natural history of human papillomavirus infections of the mucosal epithelia. APMIS. 2010;118(6-7):422-49.

3. Stanley MA. Epithelial cell responses to infection with human papillomavirus. Clin Microbiol Rev. 2012;25(2):215-22.

4. Wolf M, Garcea RL, Grigorieff N, Harrison SC. Subunit interactions in bovine papillomavirus. Proc Natl Acad Sci U S A. 2010;107(14):6298-303.

5. Cogliano V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, et al. Carcinogenicity of human papillomaviruses. Lancet Oncol. 2005;6(4):204.

6. Johnson KM, Kines RC, Roberts JN, Lowy DR, Schiller JT, Day PM. Role of heparan sulfate in attachment to and infection of the murine female genital tract by human papillomavirus. J Virol. 2009;83(5):2067-74.

7. Trottier H, Franco EL. The epidemiology of genital human papillomavirus infection. Vaccine. 2006;24 Suppl 1:S1-15.

8. zur Hausen H. Papillomavirus infections--a major cause of human cancers. Biochim Biophys Acta. 1996;1288(2):F55-78.

9. de Martel C, Plummer M, Vignat J, Franceschi S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. Int J Cancer. 2017;141(4):664-70.

10. Zhai L, Tumban E. Gardasil-9: A global survey of projected efficacy. Antiviral Res. 2016;130:101-9.

11. Molano M, Van den Brule A, Plummer M, Weiderpass E, Posso H, Arslan A, et al. Determinants of clearance of human papillomavirus infections in Colombian women with normal cytology: a population-based, 5-year follow-up study. Am J Epidemiol. 2003;158(5):486-94.

12. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. Virology. 2004;324(1):17-27.

13. Chikandiwa A, Pisa PT, Sengayi M, Singh E, Delany-Moretlwe S. Patterns and trends of HPV-related cancers other than cervix in South Africa from 1994-2013. Cancer Epidemiol. 2019;58:121-9.

14. Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med. 2003;348(6):518-27.

15. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol. 1999;189(1):12-9.

16. Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J, et al. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. J Natl Cancer Inst. 1995;87(11):796-802.

17. zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. Nat Rev Cancer. 2002;2(5):342-50.

18. Ciesielska U, Nowinska K, Podhorska-Okolow M, Dziegiel P. The role of human papillomavirus in the malignant transformation of cervix epithelial cells and the importance of vaccination against this virus. Adv Clin Exp Med. 2012;21(2):235-44.

19. Blackadar CB. Historical review of the causes of cancer. World J Clin Oncol. 2016;7(1):54-86.

20. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.

21. Williamson AL. The Interaction between Human Immunodeficiency Virus and Human Papillomaviruses in Heterosexuals in Africa. J Clin Med. 2015;4(4):579-92.

22. Mitchell SM, Sekikubo M, Biryabarema C, Byamugisha JJ, Steinberg M, Jeronimo J, et al. Factors associated with high-risk HPV positivity in a low-resource setting in sub-Saharan Africa. Am J Obstet Gynecol. 2014;210(1):81 e1-7.

23. Maria H, Dana H, Francoise M, Michael P, Jurgen W. Human papillomaviruses in Western Africa: prevalences and risk factors in Burkina Faso. Arch Gynecol Obstet. 2018;298(4):789-96.

24. Stanley M. Immune responses to human papillomavirus. Vaccine. 2006;24 Suppl 1:S16-22.

25. Kashyap N, Krishnan N, Kaur S, Ghai S. Risk Factors of Cervical Cancer: A Case-Control Study. Asia Pac J Oncol Nurs. 2019;6(3):308-14.

26. Castle PE, Giuliano AR. Chapter 4: Genital tract infections, cervical inflammation, and antioxidant nutrients--assessing their roles as human papillomavirus cofactors. J Natl Cancer Inst Monogr. 2003(31):29-34.

27. Zhao Y, Cao X, Zheng Y, Tang J, Cai W, Wang H, et al. Relationship between cervical disease and infection with human papillomavirus types 16 and 18, and herpes simplex virus 1 and 2. J Med Virol. 2012;84(12):1920-7.

28. Group FIS. Quadrivalent vaccine against human papillomavirus to prevent high-grade cervical lesions. N Engl J Med. 2007;356(19):1915-27.

29. McKeage K, Romanowski B. AS04-adjuvanted human papillomavirus (HPV) types 16 and 18 vaccine (Cervarix(R)): a review of its use in the prevention of premalignant cervical lesions and cervical cancer causally related to certain oncogenic HPV types. Drugs. 2011;71(4):465-88.

30. Descamps D, Hardt K, Spiessens B, Izurieta P, Verstraeten T, Breuer T, et al. Safety of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine for cervical cancer prevention: a pooled analysis of 11 clinical trials. Hum Vaccin. 2009;5(5):332-40.

31. Paavonen J, Naud P, Salmeron J, Wheeler CM, Chow SN, Apter D, et al. Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women. Lancet. 2009;374(9686):301-14.

32. Schiller JT, Castellsague X, Villa LL, Hildesheim A. An update of prophylactic human papillomavirus L1 virus-like particle vaccine clinical trial results. Vaccine. 2008;26 Suppl 10:K53-61.

33. Draper E, Bissett SL, Howell-Jones R, Waight P, Soldan K, Jit M, et al. A randomized, observer-blinded immunogenicity trial of Cervarix((R)) and Gardasil((R)) Human Papillomavirus vaccines in 12-15 year old girls. PLoS One. 2013;8(5):e61825.

34. Kirby T. FDA approves new upgraded Gardasil 9. Lancet Oncol. 2015;16(2):e56.

35. Harper DM, Franco EL, Wheeler C, Ferris DG, Jenkins D, Schuind A, et al. Efficacy of a bivalent L1 virus-like particle vaccine in prevention of infection with human papillomavirus types 16 and 18 in young women: a randomised controlled trial. Lancet. 2004;364(9447):1757-65.

36. Biryukov J, Meyers C. Papillomavirus Infectious Pathways: A Comparison of Systems. Viruses. 2015;7(8):4303-25.

37. Hildesheim A, Gonzalez P, Kreimer AR, Wacholder S, Schussler J, Rodriguez AC, et al. Impact of human papillomavirus (HPV) 16 and 18 vaccination on prevalent infections and rates of cervical lesions after excisional treatment. Am J Obstet Gynecol. 2016;215(2):212 e1-e15.

38. Mesher D, Panwar K, Thomas SL, Edmundson C, Choi YH, Beddows S, et al. The Impact of the National HPV Vaccination Program in England Using the Bivalent HPV Vaccine: Surveillance of Type-Specific HPV in Young Females, 2010-2016. J Infect Dis. 2018;218(6):911-21.

39. McClung NM, Gargano JW, Bennett NM, Niccolai LM, Abdullah N, Griffin MR, et al. Trends in Human Papillomavirus Vaccine Types 16 and 18 in Cervical Precancers, 2008-2014. Cancer Epidemiol Biomarkers Prev. 2019;28(3):602-9.

40. Harro CD, Pang YY, Roden RB, Hildesheim A, Wang Z, Reynolds MJ, et al. Safety and immunogenicity trial in adult volunteers of a human papillomavirus 16 L1 virus-like particle vaccine. J Natl Cancer Inst. 2001;93(4):284-92.

41. Bishop B, Dasgupta J, Klein M, Garcea RL, Christensen ND, Zhao R, et al. Crystal structures of four types of human papillomavirus L1 capsid proteins: understanding the specificity of neutralizing monoclonal antibodies. J Biol Chem. 2007;282(43):31803-11.

42. Denny L, Adewole I, Anorlu R, Dreyer G, Moodley M, Smith T, et al. Human papillomavirus prevalence and type distribution in invasive cervical cancer in sub-Saharan Africa. Int J Cancer. 2014;134(6):1389-98.

43. Li N, Franceschi S, Howell-Jones R, Snijders PJ, Clifford GM. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. Int J Cancer. 2011;128(4):927-35.

44. Ramathuba DU, Ngambi D. Knowledge and Attitudes of Women towards Human Papilloma Virus and HPV Vaccine in Thulamela Municipality of Vhembe District in Limpopo Province, South Africa. Afr J Reprod Health. 2018;22(3):111-9.

45. Delany-Moretlwe S, Kelley KF, James S, Scorgie F, Subedar H, Dlamini NR, et al. Human Papillomavirus Vaccine Introduction in South Africa: Implementation Lessons From an Evaluation of the National School-Based Vaccination Campaign. Glob Health Sci Pract. 2018;6(3):425-38.

46. Buck CB, Trus BL. The papillomavirus virion: a machine built to hide molecular Achilles' heels. Adv Exp Med Biol. 2012;726:403-22.

47. Joyce JG, Tung JS, Przysiecki CT, Cook JC, Lehman ED, Sands JA, et al. The L1 major capsid protein of human papillomavirus type 11 recombinant virus-like particles interacts with heparin and cell-surface glycosaminoglycans on human keratinocytes. J Biol Chem. 1999;274(9):5810-22.

48. Baker TS, Newcomb WW, Olson NH, Cowsert LM, Olson C, Brown JC. Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional image reconstruction. Biophys J. 1991;60(6):1445-56.

49. Klug A, Finch JT. Structure of Viruses of the Papilloma-Polyoma Type. I. Human Wart Virus. J Mol Biol. 1965;11:403-23.

50. Guan J, Bywaters SM, Brendle SA, Ashley RE, Makhov AM, Conway JF, et al. Cryoelectron Microscopy Maps of Human Papillomavirus 16 Reveal L2 Densities and Heparin Binding Site. Structure. 2017;25(2):253-63.

51. Buck CB, Day PM, Trus BL. The papillomavirus major capsid protein L1. Virology. 2013;445(1-2):169-74.

52. Lowe J, Panda D, Rose S, Jensen T, Hughes WA, Tso FY, et al. Evolutionary and structural analyses of alpha-papillomavirus capsid proteins yields novel insights into L2 structure and interaction with L1. Virol J. 2008;5:150.

53. Raff AB, Woodham AW, Raff LM, Skeate JG, Yan L, Da Silva DM, et al. The evolving field of human papillomavirus receptor research: a review of binding and entry. J Virol. 2013;87(11):6062-72.

54. Finnen RL, Erickson KD, Chen XS, Garcea RL. Interactions between papillomavirus L1 and L2 capsid proteins. J Virol. 2003;77(8):4818-26.

55. Buck CB, Cheng N, Thompson CD, Lowy DR, Steven AC, Schiller JT, et al. Arrangement of L2 within the papillomavirus capsid. J Virol. 2008;82(11):5190-7.

56. Cardone G, Moyer AL, Cheng N, Thompson CD, Dvoretzky I, Lowy DR, et al. Maturation of the human papillomavirus 16 capsid. mBio. 2014;5(4):e01104-14.

57. Kawana K, Yoshikawa H, Taketani Y, Yoshiike K, Kanda T. Common neutralization epitope in minor capsid protein L2 of human papillomavirus types 16 and 6. J Virol. 1999;73(7):6188-90.

58. Kawana Y, Kawana K, Yoshikawa H, Taketani Y, Yoshiike K, Kanda T. Human papillomavirus type 16 minor capsid protein l2 N-terminal region containing a common neutralization epitope binds to the cell surface and enters the cytoplasm. J Virol. 2001;75(5):2331-6.

59. Roden RB, Kirnbauer R, Jenson AB, Lowy DR, Schiller JT. Interaction of papillomaviruses with the cell surface. J Virol. 1994;68(11):7260-6.

60. Daniel J. Goetschius SRH, Suriyasri Subramanian, Carol Bator, Neil D. Christensen, Susan L Hafenstein. High resolution cryo EM analysis of HPV16 identifies minor structural protein L2 and describes capsid flexibility 2020.

61. Surviladze Z, Sterkand RT, Ozbun MA. Interaction of human papillomavirus type 16 particles with heparan sulfate and syndecan-1 molecules in the keratinocyte extracellular matrix plays an active role in infection. J Gen Virol. 2015;96(8):2232-41.

62. Kirnbauer R, Booy F, Cheng N, Lowy DR, Schiller JT. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. Proc Natl Acad Sci U S A. 1992;89(24):12180-4.

63. Darshan MS, Lucchi J, Harding E, Moroianu J. The l2 minor capsid protein of human papillomavirus type 16 interacts with a network of nuclear import receptors. J Virol. 2004;78(22):12179-88.

64. Angeletti PC. Seeing HPV in the New Light Offers a Glimpse of Heparin. Structure. 2017;25(2):213.

65. Sapp M, Volpers C, Muller M, Streeck RE. Organization of the major and minor capsid proteins in human papillomavirus type 33 virus-like particles. J Gen Virol. 1995;76 (Pt 9):2407-12.

66. Doorbar J, Gallimore PH. Identification of proteins encoded by the L1 and L2 open reading frames of human papillomavirus 1a. J Virol. 1987;61(9):2793-9.

67. Sapp M, Fligge C, Petzak I, Harris JR, Streeck RE. Papillomavirus assembly requires trimerization of the major capsid protein by disulfides between two highly conserved cysteines. J Virol. 1998;72(7):6186-9.

68. Ishii Y, Tanaka K, Kanda T. Mutational analysis of human papillomavirus type 16 major capsid protein L1: the cysteines affecting the intermolecular bonding and structure of L1-capsids. Virology. 2003;308(1):128-36.

69. Knappe M, Bodevin S, Selinka HC, Spillmann D, Streeck RE, Chen XS, et al. Surface-exposed amino acid residues of HPV16 L1 protein mediating interaction with cell surface heparan sulfate. J Biol Chem. 2007;282(38):27913-22.

70. Zheng ZM, Baker CC. Papillomavirus genome structure, expression, and post-transcriptional regulation. Front Biosci. 2006;11:2286-302.

71. Doorbar J, Quint W, Banks L, Bravo IG, Stoler M, Broker TR, et al. The biology and life-cycle of human papillomaviruses. Vaccine. 2012;30 Suppl 5:F55-70.

72. Berg M, Stenlund A. Functional interactions between papillomavirus E1 and E2 proteins. J Virol. 1997;71(5):3853-63.

73. Doorbar J. The E4 protein; structure, function and patterns of expression. Virology. 2013;445(1-2):80-98.

74. Doorbar J. The papillomavirus life cycle. J Clin Virol. 2005;32 Suppl 1:S7-15.

75. Pyeon D, Pearce SM, Lank SM, Ahlquist P, Lambert PF. Establishment of human papillomavirus infection requires cell cycle progression. PLoS Pathog. 2009;5(2):e1000318.

76. Munger K, Baldwin A, Edwards KM, Hayakawa H, Nguyen CL, Owens M, et al. Mechanisms of human papillomavirus-induced oncogenesis. J Virol. 2004;78(21):11451-60.

77. Schmitz M, Driesch C, Jansen L, Runnebaum IB, Durst M. Non-random integration of the HPV genome in cervical cancer. PLoS One. 2012;7(6):e39632.

78. Middleton K, Peh W, Southern S, Griffin H, Sotlar K, Nakahara T, et al. Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. J Virol. 2003;77(19):10186-201.

79. Culp TD, Budgeon LR, Christensen ND. Human papillomaviruses bind a basal extracellular matrix component secreted by keratinocytes which is distinct from a membrane-associated receptor. Virology. 2006;347(1):147-59.

80. Conway MJ, Meyers C. Replication and assembly of human papillomaviruses. J Dent Res. 2009;88(4):307-17.

81. Stubenrauch F, Laimins LA. Human papillomavirus life cycle: active and latent phases. Semin Cancer Biol. 1999;9(6):379-86.

82. Sasagawa T, Takagi H, Makinoda S. Immune responses against human papillomavirus (HPV) infection and evasion of host defense in cervical cancer. J Infect Chemother. 2012;18(6):807-15.

83. Ashrafi GH, Haghshenas MR, Marchetti B, O'Brien PM, Campo MS. E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I. Int J Cancer. 2005;113(2):276-83.

84. Chellappan S, Kraus VB, Kroger B, Munger K, Howley PM, Phelps WC, et al. Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. Proc Natl Acad Sci U S A. 1992;89(10):4549-53.

85. Cerqueira C, Schiller JT. Papillomavirus assembly: An overview and perspectives. Virus Res. 2017;231:103-7.

86. Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. Nat Rev Cancer. 2010;10(8):550-60.

87. Kines RC, Thompson CD, Lowy DR, Schiller JT, Day PM. The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding. Proc Natl Acad Sci U S A. 2009;106(48):20458-63.

88. Qi YM, Peng SW, Hengst K, Evander M, Park DS, Zhou J, et al. Epithelial cells display separate receptors for papillomavirus VLPs and for soluble L1 capsid protein. Virology. 1996;216(1):35-45.

89. Volpers C, Unckell F, Schirmacher P, Streeck RE, Sapp M. Binding and internalization of human papillomavirus type 33 virus-like particles by eukaryotic cells. J Virol. 1995;69(6):3258-64.

90. Ozbun MA. Extracellular events impacting human papillomavirus infections: Epithelial wounding to cell signaling involved in virus entry. Papillomavirus Res. 2019;7:188-92.

91. Culp TD, Budgeon LR, Marinkovich MP, Meneguzzi G, Christensen ND. Keratinocyte-secreted laminin 5 can function as a transient receptor for human papillomaviruses by binding virions and transferring them to adjacent cells. J Virol. 2006;80(18):8940-50.

92. Selinka HC, Florin L, Patel HD, Freitag K, Schmidtke M, Makarov VA, et al. Inhibition of transfer to secondary receptors by heparan sulfate-binding drug or antibody induces noninfectious uptake of human papillomavirus. J Virol. 2007;81(20):10970-80.

93. Selinka HC, Giroglou T, Nowak T, Christensen ND, Sapp M. Further evidence that papillomavirus capsids exist in two distinct conformations. J Virol. 2003;77(24):12961-7.

94. Sarrazin S, Lamanna WC, Esko JD. Heparan sulfate proteoglycans. Cold Spring Harb Perspect Biol. 2011;3(7).

95. Combita AL, Touze A, Bousarghin L, Sizaret PY, Munoz N, Coursaget P. Gene transfer using human papillomavirus pseudovirions varies according to virus genotype and requires cell surface heparan sulfate. FEMS Microbiol Lett. 2001;204(1):183-8.

96. Giroglou T, Florin L, Schafer F, Streeck RE, Sapp M. Human papillomavirus infection requires cell surface heparan sulfate. J Virol. 2001;75(3):1565-70.

97. Zybert IA, van der Ende-Metselaar H, Wilschut J, Smit JM. Functional importance of dengue virus maturation: infectious properties of immature virions. J Gen Virol. 2008;89(Pt 12):3047-51.

98. Schafer G, Blumenthal MJ, Katz AA. Interaction of human tumor viruses with host cell surface receptors and cell entry. Viruses. 2015;7(5):2592-617.

99. Day PM, Lowy DR, Schiller JT. Heparan sulfate-independent cell binding and infection with furin-precleaved papillomavirus capsids. J Virol. 2008;82(24):12565-8.

100. Shafti-Keramat S, Handisurya A, Kriehuber E, Meneguzzi G, Slupetzky K, Kirnbauer R. Different heparan sulfate proteoglycans serve as cellular receptors for human papillomaviruses. J Virol. 2003;77(24):13125-35.

101. Abban CY, Meneses PI. Usage of heparan sulfate, integrins, and FAK in HPV16 infection. Virology. 2010;403(1):1-16.

102. Dasgupta J, Bienkowska-Haba M, Ortega ME, Patel HD, Bodevin S, Spillmann D, et al. Structural basis of oligosaccharide receptor recognition by human papillomavirus. J Biol Chem. 2011;286(4):2617-24.

103. Richards KF, Bienkowska-Haba M, Dasgupta J, Chen XS, Sapp M. Multiple heparan sulfate binding site engagements are required for the infectious entry of human papillomavirus type 16. J Virol. 2013;87(21):11426-37.

104. Patterson NA, Smith JL, Ozbun MA. Human papillomavirus type 31b infection of human keratinocytes does not require heparan sulfate. J Virol. 2005;79(11):6838-47.

105. Becker M, Greune L, Schmidt MA, Schelhaas M. Extracellular Conformational Changes in the Capsid of Human Papillomaviruses Contribute to Asynchronous Uptake into Host Cells. J Virol. 2018;92(11).

106. Richards RM, Lowy DR, Schiller JT, Day PM. Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection. Proc Natl Acad Sci U S A. 2006;103(5):1522-7.

107. Cruz L, Meyers C. Differential dependence on host cell glycosaminoglycans for infection of epithelial cells by high-risk HPV types. PLoS One. 2013;8(7):e68379.

108. Surviladze Z, Dziduszko A, Ozbun MA. Essential roles for soluble virion-associated heparan sulfonated proteoglycans and growth factors in human papillomavirus infections. PLoS Pathog. 2012;8(2):e1002519.

109. Cruz L, Biryukov J, Conway MJ, Meyers C. Cleavage of the HPV16 Minor Capsid Protein L2 during Virion Morphogenesis Ablates the Requirement for Cellular Furin during De Novo Infection. Viruses. 2015;7(11):5813-30.

110. Day PM, Schiller JT. The role of furin in papillomavirus infection. Future Microbiol. 2009;4(10):1255-62.

111. Day PM, Thompson CD, Buck CB, Pang YY, Lowy DR, Schiller JT. Neutralization of human papillomavirus with monoclonal antibodies reveals different mechanisms of inhibition. J Virol. 2007;81(16):8784-92.

112. Bronnimann MP, Calton CM, Chiquette SF, Li S, Lu M, Chapman JA, et al. Furin Cleavage of L2 during Papillomavirus Infection: Minimal Dependence on Cyclophilins. J Virol. 2016;90(14):6224-34.

113. Selinka HC, Giroglou T, Sapp M. Analysis of the infectious entry pathway of human papillomavirus type 33 pseudovirions. Virology. 2002;299(2):279-87.

114. Park PW, Reizes O, Bernfield M. Cell surface heparan sulfate proteoglycans: selective regulators of ligand-receptor encounters. J Biol Chem. 2000;275(39):29923-6.

115. Shiryaev SA, Remacle AG, Ratnikov BI, Nelson NA, Savinov AY, Wei G, et al. Targeting host cell furin proprotein convertases as a therapeutic strategy against bacterial toxins and viral pathogens. J Biol Chem. 2007;282(29):20847-53.

116. Seidah NG, Mayer G, Zaid A, Rousselet E, Nassoury N, Poirier S, et al. The activation and physiological functions of the proprotein convertases. Int J Biochem Cell Biol. 2008;40(6-7):1111-25.

117. Henrich S, Cameron A, Bourenkov GP, Kiefersauer R, Huber R, Lindberg I, et al. The crystal structure of the proprotein processing proteinase furin explains its stringent specificity. Nat Struct Biol. 2003;10(7):520-6.

118. Thomas G. Furin at the cutting edge: from protein traffic to embryogenesis and disease. Nat Rev Mol Cell Biol. 2002;3(10):753-66.

119. Izaguirre G. The Proteolytic Regulation of Virus Cell Entry by Furin and Other Proprotein Convertases. Viruses. 2019;11(9).

120. Volchkov VE, Feldmann H, Volchkova VA, Klenk HD. Processing of the Ebola virus glycoprotein by the proprotein convertase furin. Proc Natl Acad Sci U S A. 1998;95(10):5762-7.

121. Zhang X, Fugere M, Day R, Kielian M. Furin processing and proteolytic activation of Semliki Forest virus. J Virol. 2003;77(5):2981-9.

122. Vey M, Schafer W, Reis B, Ohuchi R, Britt W, Garten W, et al. Proteolytic processing of human cytomegalovirus glycoprotein B (gpUL55) is mediated by the human endoprotease furin. Virology. 1995;206(1):746-9.

123. Spiesschaert B, Stephanowitz H, Krause E, Osterrieder N, Azab W. Glycoprotein B of equine herpesvirus type 1 has two recognition sites for subtilisin-like proteases that are cleaved by furin. J Gen Virol. 2016;97(5):1218-28.

124. Garten W, Klenk HD. Understanding influenza virus pathogenicity. Trends Microbiol. 1999;7(3):99-100.

125. Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG, Decroly E. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antiviral Res. 2020;176:104742.

126. Schafer G, Graham LM, Lang DM, Blumenthal MJ, Bergant Marusic M, Katz AA. Vimentin Modulates Infectious Internalization of Human Papillomavirus 16 Pseudovirions. J Virol. 2017;91(16).

127. Wang JW, Jagu S, Kwak K, Wang C, Peng S, Kirnbauer R, et al. Preparation and properties of a papillomavirus infectious intermediate and its utility for neutralization studies. Virology. 2014;449:304-16.

128. Day PM, Gambhira R, Roden RB, Lowy DR, Schiller JT. Mechanisms of human papillomavirus type 16 neutralization by l2 cross-neutralizing and l1 type-specific antibodies. J Virol. 2008;82(9):4638-46.

129. Schiller JT, Day PM, Kines RC. Current understanding of the mechanism of HPV infection. Gynecol Oncol. 2010;118(1 Suppl):S12-7.

130. Wang JW, Roden RB. L2, the minor capsid protein of papillomavirus. Virology. 2013;445(1-2):175-86.

131. Yang R, Day PM, Yutzy WHt, Lin KY, Hung CF, Roden RB. Cell surface-binding motifs of L2 that facilitate papillomavirus infection. J Virol. 2003;77(6):3531-41.

132. Cerqueira C, Samperio Ventayol P, Vogeley C, Schelhaas M. Kallikrein-8 Proteolytically Processes Human Papillomaviruses in the Extracellular Space To Facilitate Entry into Host Cells. J Virol. 2015;89(14):7038-52.

133. Bienkowska-Haba M, Patel HD, Sapp M. Target cell cyclophilins facilitate human papillomavirus type 16 infection. PLoS Pathog. 2009;5(7):e1000524.

134. Schelhaas M, Ewers H, Rajamaki ML, Day PM, Schiller JT, Helenius A. Human papillomavirus type 16 entry: retrograde cell surface transport along actin-rich protrusions. PLoS Pathog. 2008;4(9):e1000148.

135. Mikulicic S, Florin L. The endocytic trafficking pathway of oncogenic papillomaviruses. Papillomavirus Res. 2019;7:135-7.

136. Evander M, Frazer IH, Payne E, Qi YM, Hengst K, McMillan NA. Identification of the alpha6 integrin as a candidate receptor for papillomaviruses. J Virol. 1997;71(3):2449-56.

137. Fothergill T, McMillan NA. Papillomavirus virus-like particles activate the PI3-kinase pathway via alpha-6 beta-4 integrin upon binding. Virology. 2006;352(2):319-28.

138. Yoon CS, Kim KD, Park SN, Cheong SW. alpha(6) Integrin is the main receptor of human papillomavirus type 16 VLP. Biochem Biophys Res Commun. 2001;283(3):668-73.

139. Scheffer KD, Gawlitza A, Spoden GA, Zhang XA, Lambert C, Berditchevski F, et al. Tetraspanin CD151 mediates papillomavirus type 16 endocytosis. J Virol. 2013;87(6):3435-46.

140. Spoden G, Freitag K, Husmann M, Boller K, Sapp M, Lambert C, et al. Clathrin- and caveolin-independent entry of human papillomavirus type 16--involvement of tetraspanin-enriched microdomains (TEMs). PLoS One. 2008;3(10):e3313.

141. Woodham AW, Da Silva DM, Skeate JG, Raff AB, Ambroso MR, Brand HE, et al. The S100A10 subunit of the annexin A2 heterotetramer facilitates L2-mediated human papillomavirus infection. PLoS One. 2012;7(8):e43519.

142. Taylor JR, Fernandez DJ, Thornton SM, Skeate JG, Luhen KP, Da Silva DM, et al. Heterotetrameric annexin A2/S100A10 (A2t) is essential for oncogenic human papillomavirus trafficking and capsid disassembly, and protects virions from lysosomal degradation. Sci Rep. 2018;8(1):11642.

143. Schelhaas M, Shah B, Holzer M, Blattmann P, Kuhling L, Day PM, et al. Entry of human papillomavirus type 16 by actin-dependent, clathrin- and lipid raft-independent endocytosis. PLoS Pathog. 2012;8(4):e1002657.

144. Day PM, Thompson CD, Schowalter RM, Lowy DR, Schiller JT. Identification of a role for the trans-Golgi network in human papillomavirus 16 pseudovirus infection. J Virol. 2013;87(7):3862-70.

145. Day PM, Lowy DR, Schiller JT. Papillomaviruses infect cells via a clathrin-dependent pathway. Virology. 2003;307(1):1-11.

146. Smith JL, Lidke DS, Ozbun MA. Virus activated filopodia promote human papillomavirus type 31 uptake from the extracellular matrix. Virology. 2008;381(1):16-21.

147. Bergant M, Banks L. SNX17 facilitates infection with diverse papillomavirus types. J Virol. 2013;87(2):1270-3.

148. Bergant Marusic M, Ozbun MA, Campos SK, Myers MP, Banks L. Human papillomavirus L2 facilitates viral escape from late endosomes via sorting nexin 17. Traffic. 2012;13(3):455-67.

149. Popa A, Zhang W, Harrison MS, Goodner K, Kazakov T, Goodwin EC, et al. Direct binding of retromer to human papillomavirus type 16 minor capsid protein L2 mediates endosome exit during viral infection. PLoS Pathog. 2015;11(2):e1004699.

150. Bienkowska-Haba M, Williams C, Kim SM, Garcea RL, Sapp M. Cyclophilins facilitate dissociation of the human papillomavirus type 16 capsid protein L1 from the L2/DNA complex following virus entry. J Virol. 2012;86(18):9875-87.

151. Florin L, Becker KA, Lambert C, Nowak T, Sapp C, Strand D, et al. Identification of a dynein interacting domain in the papillomavirus minor capsid protein l2. J Virol. 2006;80(13):6691-6.

152. Schneider MA, Spoden GA, Florin L, Lambert C. Identification of the dynein light chains required for human papillomavirus infection. Cell Microbiol. 2011;13(1):32-46.

153. Aydin I, Weber S, Snijder B, Samperio Ventayol P, Kuhbacher A, Becker M, et al. Large scale RNAi reveals the requirement of nuclear envelope breakdown for nuclear import of human papillomaviruses. PLoS Pathog. 2014;10(5):e1004162.

154. Bordeaux J, Forte S, Harding E, Darshan MS, Klucevsek K, Moroianu J. The l2 minor capsid protein of low-risk human papillomavirus type 11 interacts with host nuclear import receptors and viral DNA. J Virol. 2006;80(16):8259-62.

155. Letian T, Tianyu Z. Cellular receptor binding and entry of human papillomavirus. Virol J. 2010;7:2.

156. Lee H, Brendle SA, Bywaters SM, Guan J, Ashley RE, Yoder JD, et al. A cryo-electron microscopy study identifies the complete H16.V5 epitope and reveals global conformational

changes initiated by binding of the neutralizing antibody fragment. J Virol. 2015;89(2):1428-38.

157. DiGiuseppe S, Bienkowska-Haba M, Guion LG, Sapp M. Cruising the cellular highways: How human papillomavirus travels from the surface to the nucleus. Virus Res. 2017;231:1-9.

158. Gambhira R, Karanam B, Jagu S, Roberts JN, Buck CB, Bossis I, et al. A protective and broadly cross-neutralizing epitope of human papillomavirus L2. J Virol. 2007;81(24):13927-31.

159. Buck CB, Pastrana DV, Lowy DR, Schiller JT. Efficient intracellular assembly of papillomaviral vectors. J Virol. 2004;78(2):751-7.

160. Buck CB, Pastrana DV, Lowy DR, Schiller JT. Generation of HPV pseudovirions using transfection and their use in neutralization assays. Methods Mol Med. 2005;119:445-62.

161. Buck CB, Thompson CD. Production of papillomavirus-based gene transfer vectors. Curr Protoc Cell Biol. 2007;Chapter 26:Unit 26 1.

162. Buck C. PD, Lowy DR, Schiller J., Thompson CD, Pang YY, Cardone G, Moyer AL,, Cheng N DI, Steven AC, Trus BL. . Production of Papillomaviral Vectors (Pseudoviruses). National Cancer Institute, Centre for Cancer Research, Oncology LoC; 2015 June 2015.

163. Shi L, Sanyal G, Ni A, Luo Z, Doshna S, Wang B, et al. Stabilization of human papillomavirus virus-like particles by non-ionic surfactants. J Pharm Sci. 2005;94(7):1538-51.

164. Schafer G, Kabanda S, van Rooyen B, Marusic MB, Banks L, Parker MI. The role of inflammation in HPV infection of the Oesophagus. BMC Cancer. 2013;13:185.

165. Steven AC, Heymann JB, Cheng N, Trus BL, Conway JF. Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. Curr Opin Struct Biol. 2005;15(2):227-36.

166. Kimanius D, Forsberg BO, Scheres SH, Lindahl E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife. 2016;5.

167. Zivanov J, Nakane T, Scheres SHW. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. IUCrJ. 2020;7(Pt 2):253-67.

168. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife. 2018;7.

169. Scheres SH. Processing of Structurally Heterogeneous Cryo-EM Data in RELION. Methods Enzymol. 2016;579:125-57.

170. Scheres SH. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol. 2012;180(3):519-30.

171. Zhang K. Gctf: Real-time CTF determination and correction. J Struct Biol. 2016;193(1):1-12.

172. Frank J. Three-dimensional Electron Microscopy of Macromolecular Assemblies. United States of Ameria: Oxford University Press; 2006.

173. Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, et al. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. J Struct Biol. 1996;116(1):190-9.

174. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605-12.

175. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods. 2017;14(4):331-2.

176. Rohou A, Grigorieff N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol. 2015;192(2):216-21.

177. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14(3):290-6.

178. Scheres SH, Chen S. Prevention of overfitting in cryo-EM structure determination. Nat Methods. 2012;9(9):853-4.

179. Rosenthal PB, Henderson R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J Mol Biol. 2003;333(4):721-45.

180. Croll TI. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr D Struct Biol. 2018;74(Pt 6):519-30.

181. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 2018;27(1):14-25.

182. Mills JE, Dean PM. Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. J Comput Aided Mol Des. 1996;10(6):607-22.

183. Zhao Q, Potter CS, Carragher B, Lander G, Sworen J, Towne V, et al. Characterization of virus-like particles in GARDASIL(R) by cryo transmission electron microscopy. Hum Vaccin Immunother. 2014;10(3):734-9.

184. Pasquato A, Dettin M, Basak A, Gambaretto R, Tonin L, Seidah NG, et al. Heparin enhances the furin cleavage of HIV-1 gp160 peptides. FEBS Lett. 2007;581(30):5807-13.

185. Bywaters SM, Brendle SA, Tossi KP, Biryukov J, Meyers C, Christensen ND. Antibody Competition Reveals Surface Location of HPV L2 Minor Capsid Protein Residues 17-36. Viruses. 2017;9(11).

186. McElwee M, Vijayakrishnan S, Rixon F, Bhella D. Structure of the herpes simplex virus portal-vertex. PLoS Biol. 2018;16(6):e2006191.

7 Appendix

Complete cell culture medium: Dulbecco's Modified Eagle Medium (DMEM) (Sigma-Alrich) supplemented with 10% heat inactivated Foetal Calf Serum (FCS) (Biochrom), penicillin (100 U/mL), streptomycin (100 U/mL)

Light CsCl: 54 g CsCl, 200 mL 1xHSB buffer, filter sterilised

Heavy CsCl: 77.6 g CsCl**,** 200 mL 1xHSB buffer, filter sterilised

7.5% SDS-PAGE separating gels: 3.8 mL 40% Acrylamide, 5 mL Tris (1.5M, pH 8.8)**,** 0.2 mL 10% Sodium Dodecyl Sulphate (SDS)**,** 11 mL dH_2O, 20 µL TEMED, 100 µL 10% ammonium persulphate

l3% SDS-PAGE stacking gel: 1.3mL 40% Acrylamide, 1.25mL Tris (1M; pH 6.8), 0.1mL 10% Sodium Dodecyl Sulphate, 7.4mL dH_2O, 20µL TEMED, 50µL ammonium persulphate (10%)

10x SDS-PAGE running buffer: 30g Solid Tris, 144g Solid Glycine, 10g SDS, made up to 1 litre with dH_2O

1x SDS-PAGE running buffer: 10x running buffer was diluted 1:9 with dH_2O to form 1X SDS-PAGE buffer

Pierce Silver Staining kit sensitizer solution: 25 µL sensitizer, 12.5 µL ultrapure dH_2O

Pierce Silver Staining kit working solution: 250 µL enhancer, 12.5mL Silver stain

Pierce Silver Staining kit staining solution: 250 µL enhancer, 12.5mL developer

2x HBS buffer: 50mM HEPES, 280 mM NaCl, 1.5 mM Na2PO4 (pH = 7.1), filter sterilised

5x HSB buffer: 125 mM HEPES, 2.5 M NaCl, 0.1% Brij58, 5 mM $MgCl_2$, 500 µM EDTA, 2.5% ethanol; filter sterilised