Structural Studies of
Non-Enveloped Viruses Associated
with Human Diseases

Daniel Luke Hurdiss

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
Astbury Centre for Structural Molecular Biology

September, 2018
Intellectual property and publication statement

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Components of this thesis contain work from jointly authored publications. Work in Chapter 3 contains figures from two jointly authored research papers published in 2016 and 2018 respectively:


The candidate prepared the virus sample, made cryo-EM grids and collected/processed the data. The candidate was also heavily involved in writing, reviewing and editing the manuscript and generating the figures.

NAR and AM supervised the project and were involved in writing, reviewing and editing the manuscript. RFT provided assistance and training for cryo-EM experiments and data processing. ELM performed all experiments relating to the virus-like particle with subsequent analysis of the cryo-EM data being performed by the candidate. ELP and MMP performed experiments to determine the infectivity of the purified virus sample.


The candidate performed all the experiments with the exception of molecular dynamics simulations. The candidate wrote the first draft and was also involved in reviewing and editing the manuscript and generating the figures.

NAR and AM supervised the project and were involved in writing, reviewing and editing the manuscript. MF performed the molecular dynamic simulations
and prepared text and figures for the corresponding section of the manuscript which was subsequently edited by the candidate. JS assisted with reviewing and editing the manuscript.

Work in Chapter 4 contains data from one jointly authored publication:


The candidate performed the cryo-EM experiments, processed the data, built and refined the model and interpreted the results. The candidate also wrote, reviewed and edited the corresponding section of the manuscript and generated the figures.

GZ provided assisted with planning and implementation of cryo-EM grid preparation under category II conditions. GZ also provided advice for model building and refinement.

In aspects of the manuscript not relating to cryo-EM: JB, HJT, and FJMvK designed research; JB, NM, RWR, JJS, HG, ALWvV, MW, KB, ED, and HJT performed research. Data analysis was performed by JB, NM, and HJT.

Reagents and analytic tools were provided by CAMdH, EdV, JMC, and RJdG. The project was supervised by EdV, RJdG, NA, TS, NAR, HJT, and FJMvK. Writing, reviewing and editing of the paper was performed by JB, GZ, NM, RWR, NA, TS, NAR, HJT and FJMvK.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Daniel Luke Hurdiss to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

© 2018 The University of Leeds and Daniel Luke Hurdiss
Acknowledgements

First and foremost, I would like to thank my ‘PhD fathers’, Neil Ranson and Andrew Macdonald, for your support, guidance and humour for the last four years. I consider you both to be not only great scientists and mentors but also good friends.

I would also like to thank Stephen Muench for introducing me to non-virus cryo-EM and for giving me the opportunity to work on such an interesting ‘side project’. I’m also extremely appreciative of Sheena Radford for her candid career advice and encouragement during the final year of my PhD.

I am immensely grateful to all my wonderful colleagues at the Astbury Centre. Firstly, I would like to thank all past and present members of the Ranson lab and EM office, in particular, the original members of the ‘banter bus’. Rebecca Thompson, thank you for showing me the EM ropes, your training, and the traumatic memories of your milky brews, will stay with me. Shaun Rawson, in spite of your profound betrayal, I thank you for all the in-house and transatlantic IT support. Emma Hesketh, thanks for being the responsible adult, and for giving me your fancy chair. Matthew Iadanza, I have much enjoyed your wicked scripts and even wickeder sense of humour. Rachel Johnson, while I don’t understand it, I respect your resolute refusal to ever drink tea despite three years of me asking. You have all been a pleasure to work and socialise with during the last four years!

I’m also very grateful to past and present members of the Macdonald group for all their help during my PhD, notably Emma Prescott, Margarita Panou and Michelle Antoni. I’d also like to thank Ethan Morgan for his contributions to my PhD project but more so for the funny conversations in the pub.

Also, I would like to thank all of the researchers I have been fortunate enough to collaborate with. From Utrecht University, I would particularly like to thank Jim Baggen, Hendrick Thibaut and Frank van Kuppeveld. Frank, your enthusiastic career support during the last year has been amazing; I look forward to working with you in the future. From Tübingen University, I would like to thank Thilo Stehle and Georg Zocher. From Umeå University I would
like to thank Niklas Arnberg and Nitesh Mistry. I would also like to thank Martin Frank from Biognos AB, Gothenburg.

I’m appreciative to all the facility staff in Leeds who have provided support during my PhD, in particular, Peiyi Wang and Martin Fuller. I would also like to thank Alistair Siebert and Dan Clare for their assistance with data collection during visits to eBIC and OPIC.

None of this work would have been possible without the generous funding of the Wellcome Trust, which I am very grateful for. In addition, I would like to express my appreciation to the Microbiology Society for provided financial support to attend the annual conferences.

I owe a massive debt of gratitude to my amazing family who have always fervently supported me in all my life’s pursuits; I thank you for everything you have ever done for me.

Finally, I would like to thank my incredible partner, Ieva Drulyte, the first and greatest discovery of my PhD.
Abstract

Non-enveloped viruses encompass many human pathogens which are responsible for a broad range of diseases that have very significant impacts on human health. By studying the structures of such viruses, insights can be gained into aspects of their lifecycle, including receptor attachment, genome uncoating and capsid assembly. This information can then serve as a structural platform for the design of targeted antivirals and vaccines. This thesis aimed to use cryo-electron microscopy (cryo-EM) to structurally characterise the structures, and receptor binding, of two important human pathogens, BK polyomavirus (BKV) and Coxsackievirus A24v (CV-A24v).

BKV causes polyomavirus-associated nephropathy and haemorrhagic cystitis in immunosuppressed patients. These are diseases for which we currently have limited treatment options. Initially, a modest resolution structure (~7 Å) of BKV is used to investigate the organisation of the viral genome and minor capsid proteins. Subsequently, high-resolution structures of BKV alone (3.8 Å) and in complex with the receptor fragment of GT1b ganglioside (3.4 Å) and heparin (3.6 Å) were determined. Collectively, these structures provide insights into capsid assembly, rationalise how GT1b enhances infection over smaller gangliosides studied previously and provide the first structural clues for glycosaminoglycan binding to BKV.

CV-A24v is responsible for large outbreaks of acute haemorrhagic conjunctivitis (AHC), a painful, contagious eye disease. Here, ICAM-1 is identified as an essential receptor for CV-A24v and the high-resolution cryo-EM structure (3.9 Å) of the virus–ICAM-1 complex is presented, which reveals the critical ICAM-1–binding residues within the capsid. These data could help identify a possible conserved mode of receptor engagement among ICAM-1–binding enteroviruses. In addition, structures of the uncoating intermediates of CV-A24v are presented which describe the molecular basis of capsid expansion. Furthermore, the molecular details of a branched pocket factor binding site in CV-A24v are described which is unique amongst currently structurally characterised human enteroviruses.
Table of Contents

Intellectual property and publication statement ........................................... ii
Acknowledgements .......................................................................................... iv
Abstract ........................................................................................................ vi
Table of Contents ............................................................................................ vii
List of Tables .................................................................................................... x
List of Figures ................................................................................................. xi
Abbreviations .................................................................................................... xiv

Chapter 1 Introduction ....................................................................................... 1
1.1 Viruses: an overview .................................................................................. 1
1.2 Structural biology of viruses ...................................................................... 4
  1.2.1 X-ray crystallography ......................................................................... 5
  1.2.2 Cryo-electron microscopy .................................................................... 7
1.3 Virus-receptor interactions ........................................................................ 8
  1.3.1 Proteinaceous receptors ..................................................................... 9
  1.3.2 Glycan receptors ............................................................................... 11
1.4 Genome packaging ...................................................................................... 12
1.5 Structural vaccinology ............................................................................... 14
1.6 Unanswered questions, aims & objectives ............................................... 16

Chapter 2 Methods and Materials ................................................................... 18
2.1 Overview of the cryo-EM workflow ......................................................... 18
  2.1.1 Sample vitrification ........................................................................... 18
  2.1.2 Screening and data collection .............................................................. 18
  2.1.3 Data pre-processing ......................................................................... 20
  2.1.4 Generating a cryo-EM map ................................................................. 20
  2.1.5 Model building and refinement ......................................................... 22
2.2 Materials ..................................................................................................... 25
  2.2.1 Chemicals and biological reagents ..................................................... 25
  2.2.2 EM materials and equipment ............................................................. 25
  2.2.3 Cells and media ................................................................................. 26
2.3 General methods ......................................................................................... 26
  2.3.1 Cell culture ....................................................................................... 26
  2.3.2 Plasmid preparation ......................................................................... 26
  2.3.3 BKV growth ..................................................................................... 27
  2.3.4 BKV purification .............................................................................. 27
Chapter 3 Structural studies of BK polyomavirus

3.1 Introduction................................................................................. 42
  3.1.1 Genome organisation and replication................................. 45
  3.1.2 Virion structure ................................................................. 46
  3.1.3 The BKV lifecycle ............................................................... 49
  3.1.4 Epidemiology and tissue tropism .......................................... 54
  3.1.5 BKV-associated diseases ..................................................... 55
  3.1.6 Current therapeutic strategies .............................................. 57
  3.1.7 Project aims ...................................................................... 59

3.2 Results part one: Intermediate-resolution studies of BKV ...... 61
  3.2.1 Production of infectious BKV .............................................. 61
  3.2.2 Production of VP1 only virus-like particles ......................... 64
  3.2.3 The sub-nanometre structure of BKV and VLP ................... 66
  3.2.4 Minor Capsid Proteins and the organisation of the dsDNA genome......................................................... 68

3.3 Discussion part one ................................................................. 72

3.4 Results part two: High-resolution studies of BKV and receptor complexes.......................................................... 74
  3.4.1 Large-scale production of BKV .......................................... 74
  3.4.2 High-resolution structures of BKV: receptor complexes ...... 75
  3.4.3 Structural comparison to SV40 .......................................... 82
  3.4.4 Location of disulphide bonds and Ca^{2+} ions ..................... 86
Chapter 3 The interaction of BKV with GT1b oligosaccharide

Chapter 3.4 The interaction of BKV with glycosaminoglycans

Chapter 3.5 Discussion part two

Chapter 4 Structural studies of Coxsackievirus A24v

Chapter 4.1 Introduction

Chapter 4.1.1 Genome organisation, translation and processing

Chapter 4.1.2 Enterovirus structure

Chapter 4.1.3 The enterovirus life cycle

Chapter 4.1.4 Epidemiology of CV-A24v

Chapter 4.1.5 Therapeutic strategies

Chapter 4.1.6 Project aims

Chapter 4.2 Results

Chapter 4.2.1 ICAM-1 Is an Essential CV-A24v Receptor

Chapter 4.2.2 The Structure of CV-A24v in Complex with ICAM-1

Chapter 4.2.3 Uncoating of CV-A24v

Chapter 4.2.4 Inhibition of CV-A24v

Chapter 4.3 Discussion

Chapter 5 Summary, future perspectives and concluding remarks

Chapter 5.1 Summary

Chapter 5.2 Continuation of the projects and future perspectives

Chapter 5.3 Concluding remarks

List of References

Appendix A Additional publications
List of Tables

Table 2-1 - Electron microscopes and detectors used during this research. ................................................................. 25
Table 2-2 – Mammalian cells used during this research .................. 26
Table 2-3 – Data collection and imaging processing parameters for BKV structures ......................................................... 36
Table 2-4 - Data collection and imaging processing parameters for CV-A24v structures ................................................... 37
Table 2-5 – Quality assessment of the BKV-GT1b atomic model...... 39
Table 2-6 - Quality assessment of CV-A24v atomic models ............ 40
Table 4-1 - Amino acid composition of CV-A24 capsid sites interacting with ICAM-1 .......................................................... 140
List of Figures

Figure 1-1 – Viruses associated with human diseases.............................. 2
Figure 1-2 – The Baltimore classification system................................. 3
Figure 1-3 – Capsid triangulation numbers........................................ 4
Figure 1-4 – Protein folds commonly observed in virus capsids............ 6
Figure 1-5 – Historical advancements in virus EM............................... 8
Figure 1-6 – Virus interactions with proteinaceous receptors.............. 10
Figure 1-7 – Glycan binding by cryo-EM and X-ray crystallography. ... 12
Figure 1-8 – Genome packaging in the Qβ capsid............................... 13
Figure 1-9 – The cryo-EM structure of rhinovirus C15a..................... 14
Figure 1-10 – Structure-guided design of stable vaccines.................. 16
Figure 2-1 – Preparation of cryo-EM grids....................................... 19
Figure 2-2 – Calculating a 3D structure from 2D projections............. 24
Figure 3-1 – Polyomaviridae.......................................................... 44
Figure 3-2 – BKV genome organisation........................................... 46
Figure 3-3 – Polyomavirus structure............................................... 48
Figure 3-4 – Ganglioside receptors of BKV........................................ 50
Figure 3-5 – BKV replication cycle................................................ 54
Figure 3-6 – Purified BKV virions are infectious................................ 63
Figure 3-7 – BK polyomavirus VLP production in HEK293TT cells..... 64
Figure 3-8 – VLPs and virions contain cellular histones and DNA....... 65
Figure 3-9 – The cryo-EM structures of BKV and VLP...................... 67
Figure 3-10 – The BKV homology model........................................ 68
Figure 3-11 – Minor capsid proteins and genome organisation........ 70
Figure 3-12 – The N-terminus of BKV VP1..................................... 71
Figure 3-13 – Large-scale production of BKV................................... 75
Figure 3-14 – Micrographs from BKV and BKV-GT1b data collection. 76
Figure 3-15 Fourier shell correlations (FSC) for BKV and BKV-GT1b. 77
Figure 3-16 – Structures of BKV and BKV-GT1b.............................. 78
Figure 3-17 – Local resolution of BKV and BKV-GT1b maps............ 79
Figure 3-18 – VP1 C-terminus mediated capsid assembly.................. 80
Figure 3-19 – Atomic model of the BKV capsid............................... 81
Figure 3-20 – Structural comparison of BKV and SV40 VP1............. 83
Figure 3-21 – The C-terminus of BKV VP1..................................... 84
Figure 3-22 – The two-fold symmetry axis of BKV............................. 85
Figure 3-23 – Location of calcium ions and disulphides in BKV
Figure 3-24 – Reduced BKV data collection and cryo-EM map
Figure 3-25 – Dissociation of BKV under reducing and chelating conditions
Figure 3-26 – Negative stain reconstruction of dissociated BKV pentons
Figure 3-27 – 3D classification of BKV
Figure 3-28 – Interaction of BKV with GT1b
Figure 3-29 – Molecular dynamics simulated interactions of BKV and GT1b
Figure 3-30 – Molecular dynamic simulations of GT1b
Figure 3-31 – BKV-heparin data collection and cryo-EM map
Figure 3-32 – Interaction of BKV with heparin
Figure 3-33 – Putative heparin density surrounding the capsid vertices
Figure 3-34 – Sequence conservation of proposed heparin binding residues
Figure 4-1 – Enteroviruses of humans
Figure 4-2 – Enterovirus genome organisation
Figure 4-3 – Enterovirus structure
Figure 4-4 – The CV-A24v sialic acid binding site
Figure 4-5 – Uncoating receptor interactions with the canyon
Figure 4-6 – Enterovirus uncoating stages
Figure 4-7 – Enteroviruses remodelling of cellular membranes
Figure 4-8 – Proposed models of enterovirus assembly
Figure 4-9 – Enterovirus replication cycle
Figure 4-10 – Global outbreaks of CV-A24v
Figure 4-11 – Identification of ICAM-1 as a CV-A24v receptor
Figure 4-12 – ICAM-1 interacts directly with CV-A24v
Figure 4-13 – ICAM-1 is essential for variant and non-variant CV-A24v
Figure 4-14 – Data collection and image processing of CV-A24v: ICAM-1
Figure 4-15 – Local resolution of the CV-A24v-ICAM-1 map
Figure 4-16 – The structure of CV-A24v in complex with ICAM-1
Figure 4-17 – Receptor footprint of ICAM-1
Figure 4-18 – Salt bridges involved in the CV-A24v–ICAM-1 interaction
Figure 4.19 – Surface electrostatics of CV-A24v and ICAM-1. .......... 138
Figure 4.20 – Hydrogen bonding interactions between ICAM-1 and CV-
A24v........................................................................................................ 139
Figure 4.21 – Conformational changes in CV-A24v and ICAM-1........ 141
Figure 4.22 – Empty CV-A24v virions in micrographs. ...................... 142
Figure 4.23 – Image processing of full and empty CV-A24v virions... 143
Figure 4.24 – Comparison of UD, SSD and FD reconstructions........ 144
Figure 4.25 – Full, altered and empty CV-A24v particles in the UD data.
.............................................................................................................. 145
Figure 4.26 – Local resolution of the full and empty CV-A24v maps.. 146
Figure 4.27 – Map quality of the full and empty CV-A24v
reconstructions....................................................................................... 147
Figure 4.28 – The CV-A24v pocket factor............................................. 148
Figure 4.29 – Global conformation changes in the CV-A24v capsid. 150
Figure 4.30 – Comparison of CV-A24v empty particle to expanded
capsids of other enteroviruses.............................................................. 151
Figure 4.31 – Collapse of the CV-A24v hydrophobic pocket............ 153
Figure 4.32 – Mechanics of the hydrophobic pocket collapse........... 154
Figure 4.33 – Disorder regions in the CV-A24v empty particle. ....... 155
Figure 4.34 – Comparison of enterovirus pocket factors.................. 157
Figure 4.35 – Comparison of inhibitor bound enteroviruses. .......... 158
Figure 4.36 – Inhibition of CV-A24v by pleconaril............................. 159
Figure 4.37 – Pleconaril binding to CV-A24v........................................ 161
Figure 4.38 – Pocket factor and pleconaril binding residues.......... 162
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAV</td>
<td>Adeno-associated virus</td>
</tr>
<tr>
<td>AFP</td>
<td>Acute flaccid paralysis</td>
</tr>
<tr>
<td>AHC</td>
<td>Acute haemorrhagic conjunctivitis</td>
</tr>
<tr>
<td>BCA</td>
<td>Bicinchonic acid</td>
</tr>
<tr>
<td>BKV</td>
<td>BK polyomavirus</td>
</tr>
<tr>
<td>CAR</td>
<td>Coxsackievirus and adenovirus receptor</td>
</tr>
<tr>
<td>CCD</td>
<td>Charge-coupled device</td>
</tr>
<tr>
<td>COPD</td>
<td>Chronic obstructive pulmonary disease</td>
</tr>
<tr>
<td>Cryo-EM</td>
<td>Cryo-electron microscopy</td>
</tr>
<tr>
<td>Cryo-ET</td>
<td>Cryo-electron tomography</td>
</tr>
<tr>
<td>CTF</td>
<td>Contrast transfer function</td>
</tr>
<tr>
<td>CV</td>
<td>Coxsackievirus (e.g. CV-A21)</td>
</tr>
<tr>
<td>DAF</td>
<td>Decay-accelerating factor</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA-binding domain</td>
</tr>
<tr>
<td>DED</td>
<td>Direct electron detector</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's modified essential medium</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>E</td>
<td>Echovirus (e.g. E-12)</td>
</tr>
<tr>
<td>eBIC</td>
<td>Electron Bio-Imaging Centre</td>
</tr>
<tr>
<td>ECL</td>
<td>Enhanced chemiluminescence</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>eGFP</td>
<td>Enhanced green fluorescent protein</td>
</tr>
<tr>
<td>EM</td>
<td>Electron microscopy</td>
</tr>
<tr>
<td>EMDB</td>
<td>Electron Microscopy Data Bank</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ERAD</td>
<td>Endoplasmic reticulum-associated protein degradation</td>
</tr>
<tr>
<td>FCV</td>
<td>Feline calicivirus</td>
</tr>
<tr>
<td>FD</td>
<td>Fully decorated</td>
</tr>
<tr>
<td>FEG</td>
<td>Field emission gun</td>
</tr>
<tr>
<td>FMDV</td>
<td>Foot and mouth disease virus</td>
</tr>
<tr>
<td>FSC</td>
<td>Fourier shell correlation</td>
</tr>
<tr>
<td>GAG</td>
<td>Glycosaminoglycan</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>HAV</td>
<td>Hepatitis A virus</td>
</tr>
<tr>
<td>HC</td>
<td>Haemorrhagic cystitis</td>
</tr>
<tr>
<td>HCE</td>
<td>Human corneal epithelial</td>
</tr>
<tr>
<td>HEK</td>
<td>Human embryonic kidney</td>
</tr>
<tr>
<td>HFMD</td>
<td>Hand, foot and mouth disease</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>HRV</td>
<td>Human rhinovirus</td>
</tr>
<tr>
<td>ICAM-1</td>
<td>Intercellular adhesion molecule-1</td>
</tr>
<tr>
<td>ICTV</td>
<td>International Committee on Taxonomy of Viruses</td>
</tr>
<tr>
<td>IgSF</td>
<td>Immunoglobulin-like superfamily</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal ribosome entry site</td>
</tr>
<tr>
<td>JAM-A</td>
<td>Junctional adhesion molecule A</td>
</tr>
<tr>
<td>JCV</td>
<td>JC polyomavirus</td>
</tr>
<tr>
<td>KTR</td>
<td>Kidney transplant recipient</td>
</tr>
<tr>
<td>LDLR</td>
<td>Low-density lipoprotein receptor</td>
</tr>
<tr>
<td>LN2</td>
<td>Liquid nitrogen</td>
</tr>
<tr>
<td>mAb</td>
<td>Monoclonal antibody</td>
</tr>
<tr>
<td>MCV</td>
<td>Merkel cell polyomavirus</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamics</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum-likelihood</td>
</tr>
<tr>
<td>NCCR</td>
<td>Non-coding control region</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>NI</td>
<td>Nucleotide/nucleoside inhibitor</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear localisation signal</td>
</tr>
<tr>
<td>NNI</td>
<td>Non- nucleotide/nucleoside inhibitor</td>
</tr>
<tr>
<td>OPIC</td>
<td>Oxford Particle Imaging Centre</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>PI4KB</td>
<td>Phosphatidylinositol 4-kinase-β</td>
</tr>
<tr>
<td>PLA2G16</td>
<td>phospholipase A2 group XVI</td>
</tr>
<tr>
<td>PSF</td>
<td>Point spread function</td>
</tr>
<tr>
<td>PtdIns4P</td>
<td>Phosphatidylinositol-4-phosphate</td>
</tr>
<tr>
<td>PV</td>
<td>Poliovirus</td>
</tr>
<tr>
<td>PVAN</td>
<td>Polymavirus associated nephropathy</td>
</tr>
<tr>
<td>PVR</td>
<td>Poliovirus receptor</td>
</tr>
<tr>
<td>PyV</td>
<td>Murine polyomavirus</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative PCR</td>
</tr>
<tr>
<td>RC</td>
<td>Replication complex</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root mean square deviation</td>
</tr>
<tr>
<td>RO</td>
<td>Replication organelle</td>
</tr>
<tr>
<td>RPTE</td>
<td>Primary renal proximal tubule epithelial</td>
</tr>
<tr>
<td>RV</td>
<td>Rhinovirus (RV-C15)</td>
</tr>
<tr>
<td>SCARB2</td>
<td>Scavenger receptor class B member 2</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering ribonucleic acid</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise ratio</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SSD</td>
<td>sub-stoichiometrically decorated</td>
</tr>
<tr>
<td>SV40</td>
<td>Simian vacuolating virus 40</td>
</tr>
<tr>
<td>SVDV</td>
<td>Swine vesicular disease virus</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-acetate-ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>TAg</td>
<td>Large T antigen</td>
</tr>
<tr>
<td>tAg</td>
<td>Small T antigen</td>
</tr>
<tr>
<td>TBSV</td>
<td>Tomato bushy stunt virus</td>
</tr>
<tr>
<td>TMV</td>
<td>Tobacco mosaic virus</td>
</tr>
<tr>
<td>truncTAg</td>
<td>Truncated T antigen</td>
</tr>
<tr>
<td>TSV</td>
<td>Trichodysplasia spinulosa-associated polyomavirus</td>
</tr>
<tr>
<td>UD</td>
<td>Undecorated</td>
</tr>
<tr>
<td>VLP</td>
<td>Virus-like particle</td>
</tr>
<tr>
<td>VPg</td>
<td>Viral protein genome-linked</td>
</tr>
<tr>
<td>XFEL</td>
<td>X-ray free-electron laser</td>
</tr>
<tr>
<td>ZIKV</td>
<td>Zika virus</td>
</tr>
</tbody>
</table>
Chapter 1 Introduction

“Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less” - Marie Curie

Our Precarious Habitat (1973) by Melvin A. Benarde

1.1 Viruses: an overview

Viruses are obligate intracellular parasites which co-opt their host cells’ replicative and translational machineries to drive the formation of progeny virions. Minimally, a virus comprises a genetic element, such as an RNA or DNA molecule, which is surrounded by a protective protein coat. In addition, viruses can be enveloped, such as influenza and human immunodeficiency virus, or non-enveloped, such as those described within this thesis. There is much morphological diversity amongst viruses. Human viruses are currently known to range in size between 20-970 nm, and may be isometric, helical or pleomorphic (Figure 1-1). Viruses are ubiquitous in nature and are found in all kingdoms of life. Viruses are responsible for a myriad of diseases which have massive impacts on human and animal health, agricultural economies and food security.

Viruses are predominantly classified according to two methods. The Baltimore classification system, which was first established in 1971, groups viruses according to whether they have an RNA or DNA genome and whether this is double-stranded or single-stranded (Baltimore, 1971) - Figure 1-2. Viruses are then further classified by the International Committee on Taxonomy of Viruses (ICTV), which defines five taxa: Order (-virales), Family (-viridae), Subfamily (-virinae), Genus (-virus) and species (King et al., 2018). Structure based classification approaches have also been proposed to complement that of the ICTV, which primarily delineates viruses into taxa based on sequence comparison (Krupovic and Bamford, 2010). The benefit of using structure-based classification is evidenced by structural similarities observed between viruses which infect different domains of life (e.g., eukaryotic adenoviruses and bacterial tectiviruses), which may represent an evolutionary relationship (Bamford, 2003).
Figure 1-1 – Viruses associated with human diseases. Size comparison of common human viruses. Genomes not shown to scale (Source: ViralZone: www.expasy.org/viralzone, SIB Swiss Institute of Bioinformatics). [published under creative commons {CC BY-NC 4.0} license].
Figure 1-2 – The Baltimore classification system. Viruses are divided into seven groups (I-VII) based on their genomic composition (Source: ViralZone: www.expasy.org/viralzone, SIB Swiss Institute of Bioinformatics). [published under creative commons {CC BY-NC 4.0} license].

It is estimated that half of all known virus families have icosahedral capsids (Zlotnick et al., 2000). A platonic icosahedron comprises twelve vertices, twenty faces and thirty edges. The simplest icosahedral virus particle, with the smallest genetic cost, can be constructed from sixty symmetry-related copies of a capsid protein, where each triangular face is formed from three of these molecules (Figure 1-3). This precise arrangement is observed in a number of viruses such as satellite tobacco necrosis virus (Larson et al., 1998).

Unfortunately, the amount of genetic material which can be encapsidated is limited by the size constraints imposed by the internal volume of the capsid. In order to overcome this limitation, many viruses employ the principle of quasi-equivalence (Caspar and Klug, 1962). Whilst an icosahedral capsid is restricted by its overall geometry, the size can be increased by incorporation of more than sixty protein subunits by the allowing them to occupy quasi-equivalent, rather than equivalent, positions in the icosahedral lattice (Figure 1-3). The ability of a virus particle to form a closed structure arises from the presence of the twelve vertices which have five-fold rotational symmetry. Therefore, to increase the dimensions of an icosahedral particle, each triangular face must comprise more protein units. This in turn increased the size of the capsid by increasing the distance between each of the five-fold vertices. Capsid triangulation can be
described by the following equation $T = h^2 + hk + k^2$, where the triangulation number ($T$) describes the distance between five-fold vertices across a hexagonal net. For example, in a simple $T=1$ virus particle, the vertices are directly adjacent to one another whereas in a $T=3$ particle, these are separated by hexamers with quasi-six-fold symmetry (Figure 1-3). The principle of quasi-equivalence describes the ability of virus coat proteins to adopt subtlety difference conformations within the icosahedral particle. This property is often mediated through flexible N and C-termini, as observed for tomato bushy stunt virus (TBSV) and simian vacuolating virus 40 (SV40) respectively (Harrison, 2017).

$$h=1, k=0 \quad t=(1)^2 + (1)(0) + (0)^2 = 1$$

$$h=1, k=1 \quad t=(1)^2 + (1)(1) + (1)^2 = 3$$

$$h=2, k=1 \quad t=(2)^2 + (2)(1) + (1)^2 = 7$$

Figure 1-3 – Capsid triangulation numbers. Principles of capsid triangulation for $T=1$, $T=3$ and $T=7$ capsids (Source: ViralZone: www.expasy.org/viralzone, SIB Swiss Institute of Bioinformatics). [published under creative commons {CC BY-NC 4.0} license].

1.2 Structural biology of viruses

A detailed understanding of virus structure provides mechanistic insights into processes such as virion assembly, genome release, receptor binding, immune recognition and inhibition by antiviral compounds. These infectious agents have been studied for decades, primarily by X-ray crystallography and with increasing frequency, cryo-electron microscopy (cryo-EM) (Rossmann, 2013). A brief
historical account of both these structural techniques and their modern advancements is included below.

1.2.1 X-ray crystallography

In 1913, William and Lawrence Bragg demonstrated that X-rays could be reflected by crystalline salts to produce diffraction patterns (Bragg and Bragg, 1913). Subsequently, in the 1920s, Astbury and Bernal performed their pioneering work on X-ray fibre diffraction, giving rise to the field of structural biology. The first X-ray fibre diffraction experiments to be performed on a virus were those of Bernal and Fankuchen in the 1930s, on tobacco mosaic virus (TMV) (Bernal and Fankuchen, 1937). Around this time, it became apparent that a number of small RNA viruses from plants and animals could diffract X-rays, indicating distinct and diverse underlying structures (Crowfoot and Schmidt, 1945; Rossmann, 2013). It wasn’t until the 1950s that Watson and Crick proposed that spherical viruses may have icosahedral symmetry (Crick and Watson, 1956; Crick and Watson, 2008). In 1958, the first protein structure, myoglobin, was determined (Kendrew et al., 1958), but it wasn’t until 1978 that Harrison and colleagues revealed the 2.9 Å resolution structure of TBSV (Harrison et al., 1978). This ground-breaking structure revealed the molecular details of the jelly-roll β-barrel (Figure 1-4A), the most commonly observed protein fold found in virus capsids (Chapman and Liljas, 2003). Furthermore, the structure of TBSV confirmed the “quasi-equivalence” theory put forward by Casper and Klug in 1962 by showing that icosahedral symmetry could be achieved by identical protein subunits making specific interactions determined by their position within the viral capsid (Casper and Klug, 1962). It wasn’t until later that a detailed structure of TMV was determined (Namba and Stubbs, 1986), which revealed that the basic building block of this helical virus particle consisted of a four-helix bundle (Figure 1-4B).
Figure 1-4 – Protein folds commonly observed in virus capsids. (A) Atomic model of the shell domain from TBSV (PDB: 2TBV), showing the jelly roll structure which comprises two four-stranded \( \beta \)-sheets which pack together to form a hydrophobic interface (Harrison et al., 1978). (B) Atomic model of the TMV coat protein (PDB: 2TMV), containing a four-helix bundle (Namba et al., 1989). Models depicted in chainbow colouring scheme.

Since the early days of X-ray crystallography, this technique has made huge contributions toward our understanding of virus structure and function. According to the Protein Data Bank (PDB), as of September 2018, over 368 icosahedral virus structures have been determined by X-ray crystallography, (Carrillo-Tripp et al., 2009), which includes many important pathogens of humans, plants and animals (Prasad et al., 1999; Silva and Rossmann, 1984; Grimes et al., 1998). Whilst X-ray crystallography remains an invaluable tool for structural virology, it is limited by the requirement to coax samples into well diffracting crystals. Generally speaking, the larger the virus particle, the more difficult it is to crystallise; however, structures of large enveloped viruses such as bacteriophages PM2 and PRD1 have been determined (Abrescia et al., 2008; Cockburn et al., 2004). In addition, whilst X-ray crystallography can produce very high-resolution structures of highly ordered virus capsids, and individual protein components, it is not useful for studying pleomorphic viruses such as influenza or metastable virus-receptor complexes. Recent advances in X-ray crystallography, such as X-ray free-electron lasers (XFELs), open up new possibilities for the technique (Grimes et al., 2018). Indeed, an XFEL approach
was recently used to collect diffraction data on bacteriophage phiX174 particles within *Escherichia coli*, demonstrating the feasibility of *in vivo* virus crystallography (Duyvesteyn et al., 2018).

### 1.2.2 Cryo-electron microscopy

Electron microscopy (EM) has been used for decades to visualise viruses, with the first electron micrograph of smallpox virus being reported in 1938 (Krause, 1938). With the introduction of negative staining in the 1950s (Hall, 1955; Brenner and Horne, 1959), EM became a routine technique for studying virus morphology (Figure 1-5A). However, it wasn’t until later that De Rosier and Klug demonstrated the potential of EM to study the three-dimensional structure of viruses (De Rosier and Klug, 1968; Crowther, Amos, et al., 1970). The next big breakthrough for EM came in the 1980s when Dubochet showed that samples could be imaged in a thin layer of vitreous ice rather than being dried and embedded in a heavy metal stain (McDowall et al., 1983). This new method preserved samples in a near native state, unlike negative staining methods which introduced significant artefacts. Many of the first samples to be imaged by cryo-electron microscopy were viruses, including Semliki forest virus, bacteriophage T4 and SV40 (Adrian et al., 1984; Dubochet et al., 1986). Subsequently, very many virus structures were determined by cryo-EM, although the achievable resolution was typically limited to ~10 Å. Whilst much biological insight was gained, X-ray crystallography remained the only viable method for characterising viruses at high-resolution (Haguenau et al., 2003; Rossmann, 2013). Advancements in low-dose image acquisition (I.A.M. Kuo and Glaeser, 1975) and image processing methods (e.g., particle averaging (Frank, 1975)) eventually led to the first sub-nanometre cryo-EM reconstructions of viruses, allowing visualisation of α-helices in the hepatitis B virus capsid (Conway et al., 1997) and TMV (Jeng et al., 1989). In 2008, high-resolution cryo-EM structures of rotavirus, bacteriophage ε15 and cytoplasmic polyhedrosis virus were published (X. Zhang et al., 2008; W. Jiang et al., 2008; Yu et al., 2008). This represented a significant milestone in structural virology as for the first time, a protein backbone could be traced independently of X-ray crystallography.

Cryo-EM has become a mainstream tool to determine high-resolution virus structures, providing the sample is of good biochemical quality (Kaelber et al., 2017). Improved microscope stability, direct-electron detectors and better image
processing algorithms have all contributed to this so called ‘resolution revolution’ (Kühlbrandt, 2014). As with all structural techniques, cryo-EM has its advantages and disadvantages. The benefit of cryo-EM is that there is no requirement for protein crystals, which is often the limiting step in X-ray crystallography. Unfortunately, cryo-EM does not yet have the high-throughput capability of X-ray crystallography and in general, the average resolution achievable (3-4 Å) is poorer (Ho and Reddy, 2018). However, this paradigm may be set to change as the first sub-2 Å resolution cryo-EM structure of a virus was recently reported - Figure 1-5B-C (Y.Z. Tan et al., 2018).

**Figure 1-5 – Historical advancements in virus EM.** (A) Electron micrograph of negatively stained Turnip Yellow Mosaic Virus particles. From (Brenner and Horne, 1959), Biochem. Biophys. Acta., 34, p103. (B) Cryo-EM reconstruction of Adeno-associated virus (AAV) 2 (EMDB: 9012) determined to ~1.86 Å resolution, coloured according to the radial colouring scheme (Å). Representative density from (B) containing the corresponding atomic coordinates (PDB: 6E9D) (Y.Z. Tan et al., 2018).

### 1.3 Virus-receptor interactions

To initiate an infection, viruses must first gain access to the cell. The first step in this process is engagement of one or more receptor molecules on the cell surface, which can be proteins (e.g., integrins) or sugars (e.g., sialylated glycans). The choice of receptor molecule often dictates a virus’s tropism and changes in receptor usage can lead to emergence of new strains with altered pathogenicity or host range (Bhella, 2015). In addition, a detailed understanding of receptor recognition can provide a framework for the development of inhibitors which interfere with the process of cellular entry (Zhou and G. Simmons, 2012). Therefore, structural studies to understand the molecular basis of receptor
specificity are of great importance. For many years, cryo-EM and X-ray crystallography have been used in combination to investigate how viruses recognise their cognate receptor molecules (Rossmann, 2013). A summary of such studies in relation to non-enveloped viruses is included below.

1.3.1 Proteinaceous receptors

The study of non-enveloped viruses’ interactions with proteinaceous receptors has been primarily driven by cryo-EM (Rossmann et al., 2002). The majority of such studies have focused on picornaviruses and their interactions with members of the immunoglobulin-like superfamily (IgSF) of proteins such as integrins and cell-adhesion molecules. The archetype example is rhinovirus and its proteinaceous receptor intercellular adhesion molecule-1 (ICAM-1) (Kolatkar et al., 1999). Subsequent studies on ICAM-1 binding to Coxsackievirus A21 (CV-A21) revealed a similar mechanism of receptor recognition (Figure 1-6A) (Xiao et al., 2001; Xiao et al., 2005). These structures demonstrated that slender IgSF molecules are able to bind in narrow surface depressions on the virus surface which are less accessible to neutralizing antibodies. Thus, residues which participate in receptor engagement can remain largely conserved whilst allowing residues accessible to neutralising antibodies to mutate rapidly (Rossmann et al., 2002). Other examples of picornavirus-receptor interactions studied by cryo-EM include Coxsackievirus B3 (CV-B3) and Coxsackievirus and adenovirus receptor (CAR) (Yoder et al., 2012) (Figure 1-6B) and echovirus 12 (E-12) in complex with decay-accelerating factor (DAF) (Pettigrew et al., 2006) (Figure 1-6C). In addition to studies of picornaviruses, the interaction of feline calicivirus (FCV) and its proteinaceous receptor junctional adhesion molecule A (JAM-A) (Figure 1-6D) was also determined by cryo-EM (Bhella and Goodfellow, 2011). It is worth noting that X-ray crystallography has been used to successfully determine the structure of human rhinovirus 2 in complex with the calcium containing module of the very low-density lipoprotein receptor (Verdaguer et al., 2004). In this instance, a small fragment of the receptor could be utilised, thus making the complex amenable to crystallisation.

Whilst cryo-EM has been a valuable tool for studying virus interactions with proteinaceous receptors, the resolution typically achievable for such structures has typically been limited to ~8-20 Å (Bhella, 2015). At this resolution, interpretation of binding interfaces is typically limited to rigid-body fitting of crystal
structures into EM density maps. Recently, complexes of foot and mouth disease virus (FMDV) and αvβ6 integrin were determined to global resolutions of 3.1 Å and 3.5 Å. However, as a result of multiple possible binding modes, the receptor density was blurred and therefore not readily interpretable. Using an innovative localised reconstruction method, two possible modes of αvβ6 integrin binding were delineated at resolutions between 8.6-12.3 Å (Kotecha et al., 2017). Nevertheless, a sub-4 Å resolution cryo-EM structure of a virus-receptor complex with readily interpretable side-chain information at the interfacing region has yet to be determined.

Figure 1-6 – Virus interactions with proteinaceous receptors. Quasi-atomic models of virus-receptor complexes derived from cryo-EM maps. (A) CV-A21 in complex with ICAM-1 (PDB: 1Z7Z). (B) CV-B3 in complex with CAR (PDB: 1JEW). (C) E-12 in complex with DAF (PDB: 1UPN). (D) FCV in complex with JAM-A (EMDB: 1946). Virus capsids and receptor molecules are coloured blue and pink respectively. Adapted from (Bhella, 2015) [published under creative commons {CC BY 4.0} license].
1.3.2 Glycan receptors

In contrast to studies of proteinaceous receptors, virus-glycan interactions have been largely characterised by X-ray crystallography (Blaum et al., 2018). Glycan receptors such as glycosaminoglycans (GAGs) and sialylated oligosaccharides are conformationally flexible and charged molecules which typically offer few possibilities for hydrophobic contact. These molecules usually bind shallow, surface exposed regions on the virus surface and have dissociation constants in the millimolar range (Neu et al., 2011). These properties make them well suited to crystallographic soaking experiments where glycans of interest can be soaked into virus crystals. A number of virus-glycan interactions have been determined in this way, such as FMDV in complex with heparin (Fry et al., 2005), and enterovirus D68 (EV-D68) in complex with sialic acid. (Yue Liu, Sheng, Baggen, et al., 2015)

However, in some cases X-ray crystallography is not suitable for studying glycan binding, either because the virus cannot be produced in sufficient quantities for crystallography or because glycan binding interferes with the crystal lattice (Lerch and Chapman, 2012). In such cases cryo-EM has been used to shed light on these interacts, with typically modest resolution. Early attempts to study glycan binding by cryo-EM were performed on AAV in order to identify the GAG binding site. By generating difference maps between bound and unbound reconstructions, the site of heparin binding was localised to the shoulder region of the spikes surrounding the three-fold symmetry axis (O'Donnell et al., 2009). More recently, structures of AAV bound to the GAG analogue sucrose octasulfate (Figure 1-7) and fondaparinux were determined at 4.8 Å and 2.8 Å resolution respectively (Xie et al., 2013; Xie et al., 2017). In both cases the ligand density was resolved at ~5 Å resolution, sufficient to identify the approximate site of binding. In addition, cryo-EM was recently used to visualise heparin binding to human papillomavirus 16 at ~4.3 Å resolution (J. Guan et al., 2017). Whilst these studies demonstrate the increasing utility of cryo-EM for studying virus-glycan binding, the current resolution range of ~4.5-5 Å is significantly worse than that of x-ray crystallography, which can produce sub-2 Å resolution in favourable cases (Figure 1-7).
1.4 Genome packaging

Mechanisms of genome packaging differ greatly between viruses with different genomic compositions. For example, the genome in portal-containing dsDNA viruses, such as herpesviruses and T7 bacteriophage, are highly compacted into a solenoid structure, and this compaction is driven by powerful packaging motors (D.W. Bauer et al., 2015; Cerritelli et al., 1997). These highly pressurised genomes do not typically make sequence specific interactions with the viral capsid. In contrast, the genomes of ssRNA viruses collapse during encapsidation and are able to form sequence specific RNA secondary structures (Speir et al., 1995; Gopal et al., 2012). For some ssRNA viruses, these structured regions of the genome, which typically form stem-loop structures, are proposed to act as ‘packaging signals’ which mediate selective encapsidation of the viral genome through sequence-specific interactions with the viral capsid proteins (Stockley et al., 2013). Regions of ssRNA genomes that interact with the viral capsid, and which adhere to its icosahedral symmetry, have been successfully resolved by X-ray crystallography and cryo-EM (Schneemann, 2006; Kaelber et al., 2017). For example, the crystal structure of pariacoto virus revealed ~25 ordered RNA bases situated beneath the icosahedral two-fold axes which were packaged as an A-type duplex (L. Tang et al., 2001). For packaged genomes which do not adhere to the icosahedral symmetry of the capsid, structural information is lost.
during symmetry averaging. The most stunning structural examples of asymmetrically packaged viral genomes come from cryo-EM studies of bacteriophages MS2 and Qβ (Koning et al., 2016; Dai et al., 2016; Gorzelnik et al., 2016). These studies demonstrated that a single copy of the maturation protein replaces one of the capsid protein dimers, thus breaking the icosahedral symmetry of the capsid. Remarkably, the asymmetry which results from binding of the maturation protein is observed across the entire virion, resulting in a single unique structure with well-defined density for almost the entire packaged genome (Figure 1-8).

**Figure 1-8 – Genome packaging in the Qβ capsid.** (A) Cutaway view of the asymmetric cryo-EM reconstruction of Qβ, showing the maturation protein (pink), viral genome (yellow) and capsid (blue, green and red). A large RNA helix (red) spans the two ends of the equator which are indicated by a purple and an orange circle. Left (B) and right (C) outside views of the Qβ map are shown with the equator indicated by a dashed black line (Figure taken from (Gorzelnik et al., 2016)).
1.5 Structural vaccinology

Structural vaccinology encompasses the use of structural information to inform the rational design of immunogens (Cozzi et al., 2013). As a result of the rapid advancements in cryo-EM, it is now possible to rapidly determine high-resolution structures of viral pathogens from a relatively small amount of sample (Earl and Subramaniam, 2016). These structures provide molecular details of viral surface features which may be targeted by antibodies or antivirals. In February 2016, the World Health Organisation declared Zika virus (ZIKV) ‘a public health emergency of international concern’, following its rapid spread in the Americas (World Health Organisation, 2016). The following month, the 3.8 Å resolution structure of ZIKV was published, which revealed the molecular details of a glycosylation site believed to be a critical for host cell attachment (Sirohi et al., 2016). Furthermore, the long sought-after structure of a rhinovirus C (RV-C) species member was also reported in 2016 (Yue Liu et al., 2016). This high-resolution cryo-EM reconstruction revealed novel immunogenic features, including spike-like protrusions on the virus surface and a putative sialic acid binding site (Figure 1-9).

![Figure 1-9](image)

**Figure 1-9 – The cryo-EM structure of rhinovirus C15a.** An isosurface representation of the 2.8 Å resolution cryo-EM structure of RV-C15a (left). An enlarged view of a single asymmetric unit is shown (purple) with a putative sialic acid binding site coloured orange (Figure taken from (Earl and Subramaniam, 2016)).
Cryo-EM has proven to be a useful tool for studying the molecular basis of antibody mediated neutralisation of viruses. As with complexes of proteinaceous receptors, these reconstructions have typically been limited to low resolution, however, a number of sub-4 Å maps have now been reported (Lawson et al., 2011). These immune complexes can provide insights into different mechanisms of virus neutralisation. For example, a study by Dong et al demonstrated that antibody binding neutralises rhinovirus B14 (RV-B14) by inducing conformational changes which result in genome release (Dong et al., 2017). Structural studies of virus-antibody complexes can also permit detailed epitope mapping, which can help identify promising vaccine candidates. Such a study was reported on Coxsackievirus A6 (CV-A6), a major cause of hand, foot and mouth disease (HFMD) in humans. The authors determined the high-resolution structure of an infectious particle of CV-A6 in complex with a neutralising antibody, which allowed them to pinpoint the location of an immune dominant neutralising epitope. Furthermore, this epitope was also shown to be present on a non-infectious procapsid of CV-A6, thus demonstrating its potential as a vaccine candidate (Xu et al., 2017).

High-resolution virus structures, regardless of how they are determined, allow the rational design of capsids with improved antigenic stability. One example of this approach comes from studies of FMDV, a highly infectious virus which causes disease in cattle, sheep and pigs. By examining the structures of different FMDV serotypes, a histidine residue at the interface between adjacent protein subunits was selected for mutagenesis to a cysteine. The resulting engineered capsid displayed improved stability at high temperature and low pH as a result of a disulphide crosslink across the two-fold symmetry axis (Figure 1-10). Importantly, the engineered FMDV capsids were able to elicit a prolonged protective immune response in vaccinated animals (Porta et al., 2013). In addition, cryo-EM was recently used to determine the structure of genetically-stabilised, poliovirus-like particles expressed in plants. The results demonstrated that these particles were morphologically indistinguishable from the wild-type virus (Marsian et al., 2017).
Figure 1-10 – Structure-guided design of stable vaccines. (A) Crystal structure of FMDV serotype A22 with VP1, VP2 and VP3 coloured blue, green and red respectively. (B) Enlarged view of two adjacent capsid pentamers. (C) Histidine residue located at the two-fold symmetry axis which was selected for mutagenesis. (D) Electron density for the wild-type and (E) disulphide crosslinked capsid (Figure adapted from (Porta et al., 2013) [published under creative commons {CC BY 4.0} license]).

1.6 Unanswered questions, aims & objectives

Whilst cryo-EM has enabled many clinically important human viruses to be structurally characterised, many more remain to be studied. This thesis will focus on two such non-enveloped human pathogens, BK polyomavirus (BKV) and Coxsackievirus A24v (CV-A24v). For each, the lack of therapeutic strategies to treat the diseases they are responsible for presents an urgent unmet need.

BKV is an opportunistic pathogen responsible for a number of diseases in immunosuppressed individuals, such as haemorrhagic cystitis (HC) and polyomavirus-associated nephropathy (PVAN) (Bennett et al., 2012). Despite
their clinical importance, polyomaviruses are underrepresented in the EMDB. Excluding data presented in this thesis, only seven maps are available with resolutions ranging between 11-35 Å (Lawson et al., 2011). This work aimed to address this lack by determining high-resolution structures of BKV, and by doing so, provide insights into genome packaging, capsid assembly and its interactions with cellular receptors. Furthermore, there is increasing interest in the possibility that polyomavirus-associated diseases could be treated with monoclonal antibodies (mAbs) (Jelcic et al., 2015). The availability of high-resolution capsid structures of human polyomaviruses, such as BKV, represents an important first step towards characterising where these potentially therapeutic mAbs bind.

In contrast to the opportunistic nature of BKV, CV-A24v is responsible for pandemic outbreaks of acute haemorrhagic conjunctivitis (AHC), a painful and highly contagious infection of the eye. CV-A24v can also cause neurological symptoms such as acute flaccid paralysis in a small number of cases (Pons-Salort et al., 2015). Whilst the high-resolution structure of CV-A24v has been determined by X-ray crystallography (Zocher et al., 2014), a number of important questions remain, such as:

1. What is the uncoating receptor for CV-A24v and how does this interact with the virus capsid?
2. What do the uncoating intermediates of CV-A24v look like?
3. Can CV-A24v be inhibited with capsid binding compounds?

This thesis sought to address the structural aspects of these questions and by doing so provide insights into the CV-A24v lifecycle. Most importantly, this work aims to identify potential therapeutic strategies to ease clinical symptoms, or potentially limit the spread of CV-A24v during large outbreaks of AHC.

In more general terms, this thesis will explore the utility of cryo-EM to study virus interactions with protein and glycan receptors at high-resolution. Prior to this project, no cryo-EM maps of either receptor type were resolved at better than 4 Å resolution. The ability to produce such structures represents a significant boon for the field of structural virology, allowing these important interactions to be dissected in greater detail and thus providing novel biological insights into early events in virus entry.
Chapter 2 Methods and Materials

2.1 Overview of the cryo-EM workflow

The primary technique used to generate the structures described in this thesis is cryo-EM, and as such a brief summary of the steps involved in going from sample to 3D structure is included below. For a more in-depth overview of sample preparation, image formation and image processing, a number of comprehensive reviews are available (Drulyte et al., 2018; Thompson et al., 2016; Y. Cheng et al., 2015; Milne et al., 2013; Orlova and Saibil, 2011).

2.1.1 Sample vitrification

Once a purified sample is obtained, it must be prepared for imaging in the transmission electron microscope, which is maintained under high vacuum. In order to protect the sample from dehydration and radiation damage, it must be cryo-immobilised in a layer of vitreous ice, which preserves the specimen in an aqueous environment (L.A. Baker and J.L. Rubinstein, 2010). Purified sample is first applied to a glow discharged EM grid. To maximise sample contrast during imaging in the electron microscope, the sample must be suspended in the thinnest layer of ice possible whilst still being fully embedded. The most conventional way to generate a thin film of sample is by blotting away excess solvent present on the EM grid using filter paper. Finally, the sample needs to be cryogenically immobilised. In order to achieve the extremely quick freezing required for vitrification (~10^6 °C/s), the sample is plunged into a cryogen, such as liquid ethane or propane, which is maintained at -178 °C by liquid nitrogen (Dubochet et al., 1988).

2.1.2 Screening and data collection

To identify regions of cryo-EM grids with suitable ice thickness, particle distribution and minimal damage, incremental screening of grid squares is performed. This is carried out at lower magnification and with a minimal electron dose to prevent extensive pre-irradiation of the biological sample (Murata and Wolf, 2018). Once an area suitable for data collection has been identified, the electron beam can be focussed at a set angle and distance away from desired imaging area. Improvements in software development now mean that it is
possible to collect data in a semi-automatic fashion. A number of programmes are available to do this including EPU, Leginon (Carragher et al., 2000) and SerialEM (Mastronarde, 2005), all of which follow the same general principles. Briefly, a number of low magnification images are acquired and computational stitched together to produce a mosaic pattern of the entire grid. From here the user can select which regions they want to collect data from. Data collection then proceeds autonomously, maximising the number of images which can be acquired in the available time slot.

Figure 2-1 – Preparation of cryo-EM grids. Firstly, the sample is applied to a glow-discharged (hydrophilic) holey or lacy carbon cryo-EM grid. Excess solvent is then removed by blotting with filter paper for a user-defined amount of time. The blotted grid is subsequently plunged into liquid ethane or propane that has been precooled at liquid nitrogen temperature. This results in the biomolecule of interest being embedded in a thin film of amorphous ice. (Figure taken from (Murata and Wolf, 2018) [published under creative commons {CC BY 4.0} license]).
2.1.3 Data pre-processing

Before collected images can be used to determine a high-resolution structure, several pre-processing steps must be performed. Modern cryo-EM images are acquired using direct electron detectors (DEDs) which have a greatly improved quantum efficiency compared to previously used charge-coupled devices (CCDs) (McMullan et al., 2014). The high frame rate of DEDs means that images can be collected as a series of frames or a ‘movie’. Using these frames, motion correction software such as Unblur (Grant and Grigorieff, 2015) and Motioncorr (Zheng et al., 2017) can compensate for the blurring effect of stage drift and beam-induced movement (Brilot et al., 2012). Concurrently with motion correction, dose weighting can be performed to mitigate the effect of radiation damage by weighting individual frames based on their accumulated electron dose (Grant and Grigorieff, 2015).

Biological specimens are weak phase objects, meaning they produce minimal amplitude contrast and weak phase shift which results in very little phase contrast (Orlova and Saibil, 2011). To improve contrast between the biological specimen of interest and the surrounding solvent, phase contrast can be increased during image acquisition by collecting data with a range of defocus values. The resulting images are modulated in frequency space by the contrast transfer function (CTF). The CTF mathematically describes the point spread function (PSF) of the microscope. This arises from the cumulative effect of imperfections within the imaging setup, such as aberrations within the lens systems. Therefore, recorded images are a convolution between the CTF of the microscope, information from the imaged object and noise resulting from the imperfect nature of detectors. The CTF for each image must be estimated so that it can be deconvoluted later. There are currently a number of programmes available to estimate CTF values (e.g. ctffind (Rohou and Grigorieff, 2015), gctf (K. Zhang, 2016) and e2ctf.py (Bell et al., 2016)).

2.1.4 Generating a cryo-EM map

Once motion correction, dose weighting and CTF estimation have been performed, the position of particles within each micrograph must be computationally defined. The coordinates for particles can be selected manually allowing greater control over which particles are chosen, however this is a slow
and laborious process, particularly for large cryo-EM data sets. Alternatively, a number of automatic correlation-based particle picking programmes are now available such as Relion-autopick (Scheres, 2012), e2boxer.py (G. Tang et al., 2007), Scipion (la Rosa-Trevin et al., 2016) and FindEM (Roseman, 2004).

Individual particles extracted from cryo-EM images have a low signal-to-noise ratio (SNR). The SNR can be improved by performing 2D classification, where particles with similar orientations are aligned and averaged together. Classification can be performed using user-defined references or in a reference-free manner, which helps limit user bias. A number of approaches can be used to perform 2D class averaging including ‘K means classification’, ‘equal distributed K means’ and ‘maximum likelihood (ML)’ (Sigworth, 2016). In the ML approach, each particle contributes to all classes and is assigned a probability weighting. By weighting particles based on the likelihood they belong to each class, the ML method can sort out ‘junk particles’, which often align equally well or poorly with a number of classes. Classification can also be performed in three dimensions. The programme Relion (Scheres, 2012) implements a ML 3D classification approach where the user selects the number of desired classes and then particles are randomly assigned to these based upon an initial model, leading to subtle variations in the classes. This process is then iteratively repeated until the classification stabilises. 3D classification is useful for identifying different structural states and removing ‘junk particles’ which may not have been apparent from 2D classification.

Once a homogeneous subset of particles has been obtained, a final 3D reconstruction step is performed. The principle behind obtaining a 3D structure from 2D projections is described by the central section theorem (Crowther, DeRosier, et al., 1970). This states that the 2D Fourier transform of a projection corresponds to a central slice through the same 3D object. Thus, by determining the angular relationship between 2D projections, the structure of the 3D object can be calculated (Figure 2-2). A typical approach to this is by projection matching, where an initial starting model is back projected and then compared to the experimental data. This is used to bootstrap a new model and the process is repeated in an iterative fashion until the model reflects the experimental data. As with classification, this process of angular assignment can be performed in a ML manner, which helps to account for uncertainty. In order to recover high-
resolution information during this process, the CTF values determined earlier are used to correct for the PSF of the defocused image (Mindell and Grigorieff, 2003). A number of popular image processing software packages available for determining cryo-EM structures, such as Relion (Scheres, 2012), cryoSPARC (Punjani et al., 2017) and Frealign (Grigorieff, 2016).

In modern cryo-EM software, data is refined as two independent half-maps, each typically containing either odd or even particles. The quality of the 3D EM density map is determined by the Fourier Shell Correlation (FSC) which measures the correlation between the two half maps in Fourier space. At increasing resolution, poorer agreement is observed between half-maps, and FSC values of 0.5 (van Heel and Schatz, 2005) and 0.143 (Rosenthal and Henderson, 2003) have been proposed as thresholds to indicate the resolution cut-off of the reconstruction. The map obtained from 3D refinement is often less featured than would be expected from the estimated resolution. This results from imperfections in the imaging system which results in the loss of information at high-spatial frequencies as an envelope function. To recover these details, a process of map or B-factor sharpening is performed where high-resolution signal is boosted (Rosenthal and Henderson, 2003).

2.1.5 Model building and refinement

For high-resolution cryo-EM reconstructions, typically better than 4 Å, it is possible to build an atomic model. Depending on quality of the map, this can be performed de novo or by rigid-body fitting an existing model into the EM density to serve as a starting point. The typical workflow is similar to that of crystallography where the model is edited manually in programs such as Coot (Emsley and Cowtan, 2004), and then subject to automated refinement in Phenix (Echols et al., 2014) or refmac (Murshudov et al., 2011). This process is repeated iteratively until the atom-model geometry is optimised. The quality of the refined atomic coordinates can be assessed using crystallographic validation tools such as MolProbity (V.B. Chen et al., 2010) which generate a number of statistics and check for stereochemical outliers within the model.
For lower resolution cryo-EM maps (between 4–9 Å), atomic modelling is not possible but secondary structure features are often resolved. This permits existing models to be flexibly fitted into the EM density. A number of programmes are available to do this including MDFF (Trabuco et al., 2011), FlexEM (Topf et al., 2008) and Rosetta (DiMaio et al., 2015), each of which work in a subtly different fashion. One approach to flexibly fitting atomic coordinates into an intermediate resolution cryo-EM map is to perform a molecular dynamics simulation on an atomic model of interest which is fitted as a rigid body to the corresponding EM map. During the simulation the atomic coordinates are encouraged to move into the experimental map by allowing the model to ‘relax’ according to a weighting applied to the EM density. Secondary structure restraints can be applied to an atomic model during this process in order to preserve proper stereochemistry. One drawback of modelling into low-resolution EM maps is that such experiments are prone to overfitting and thus great care must be taken when interpreting the results.
Figure 2-2 – Calculating a 3D structure from 2D projections. In the cryo-EM grid, particles are randomly oriented throughout the vitreous ice layer. The recorded images are projections of different views of the sample and contain containing noise and distortion resulting from the microscope CTF. The Fourier transform of these 2D projections correspond to central sections through the 3D Fourier transform. By calculating the angular relationship between 2D projections, the 3D Fourier transform can be generated. A 3D reconstruction of the sample is obtained by performing a reverse Fourier transformation. (Figure taken from (Ranson and Stockley, 2011) with authors permission.)
2.2 Materials

2.2.1 Chemicals and biological reagents

Unless stated otherwise, all chemicals and reagents were purchased from Sigma-Aldrich or Fisher Scientific Ltd. All restriction enzymes and ligases were purchased from New England Biolabs (NEB).

2.2.2 EM materials and equipment

Continuous carbon lacy grids were purchased from Agar Scientific and all Quantifoil grids purchased from Quantifoil. Grids were glow discharged using either a Cressington 208 or a Pelco EasiGlow immediately prior to use. Cryo grids were prepared using either Vitrabot IV (FEI) at nominal 100% humidity and 4°C or Leica EM GP at nominal 100% humidity and 8°C. All cryo-EM grids were stored in storage dewars under liquid nitrogen (LN₂).

Transmission electron microscopes used in this research are outlined in Table 2-1.

Table 2-1 - Electron microscopes and detectors used during this research.

<table>
<thead>
<tr>
<th>Microscope</th>
<th>Location</th>
<th>Voltage (keV)</th>
<th>Electron Source</th>
<th>Detector</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEI Tecnai T12</td>
<td>University of Leeds</td>
<td>120</td>
<td>LaB6</td>
<td>Gatan US1000XP CCD</td>
</tr>
<tr>
<td>FEI Tecnai F20</td>
<td>University of Leeds</td>
<td>200</td>
<td>FEG</td>
<td>Gatan US4000/SP CCD &amp; K2 Summit DED</td>
</tr>
<tr>
<td>FEI Polara</td>
<td>University of Oxford</td>
<td>300</td>
<td>FEG</td>
<td>Gatan K2 Summit DED</td>
</tr>
<tr>
<td>FEI Titan Krios I</td>
<td>University of Leeds</td>
<td>300</td>
<td>FEG</td>
<td>Falcon II DED</td>
</tr>
<tr>
<td>FEI Titan Krios II</td>
<td>University of Leeds</td>
<td>300</td>
<td>FEG</td>
<td>Gatan K2 Summit DED</td>
</tr>
</tbody>
</table>
2.2.3 Cells and media

Mammalian cells used in this research are outlined in Table 2-2.

**Table 2-2 – Mammalian cells used during this research.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Supplier and code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vero</td>
<td>Cell line derived from African green monkey kidney epithelial cells</td>
<td>ATCC® CCL-81™</td>
</tr>
<tr>
<td>RPTE</td>
<td>Primary renal proximal tubule epithelial cells</td>
<td>Lonza CC-2553</td>
</tr>
<tr>
<td>HEK293TT</td>
<td>Cell line derived from human embryonic kidney cells, expressing SV40 small and large T antigens</td>
<td>Dr Chris Buck, National Cancer Institute, Bethesda, USA</td>
</tr>
</tbody>
</table>

2.3 General methods

2.3.1 Cell culture

Vero cells were maintained in Dulbecco’s modified essential medium (DMEM) with 10% foetal bovine serum and 50 U/ml penicillin and streptomycin. Primary RPTE cells (Lonza) were maintained in renal epithelial growth medium with the REGM BulletKit™ supplements (Lonza). 293TT cells (Buck et al., 2004) were cultured in DMEM with 10% foetal bovine serum (FBS) and non-essential amino acids (Life Technologies). All cells were grown at 37°C with 5% CO₂ in a humidified incubator.

2.3.2 Plasmid preparation

The pGEM7 plasmid containing the BKV Dunlop genome (a kind gift from Professor Michael Imperiale, University of Michigan) was prepared using a QIAGEN® Plasmid Maxi Kit according to the manufacturer’s guidelines. The DNA concentration was determined using a NanoDrop spectrophotometer (Thermo Fisher Scientific). A 1 μl aliquot of the DNA solution was applied to the sample pedestal and the DNA concentration (ng/μl) was determined from the A₂₆₀ reading.
2.3.3 BKV growth

The BKV Dunlop genome was excised from the pGEM7 Dunlop plasmid by BamHI (NEB) digestion and then recircularised using T4 Ligase (NEB). Vero cells were seeded into T75 flasks and transfected with 4 µg of DNA using NanoJuice™ (Novagen) according to the manufacturer’s instructions with a DNA to Core ratio of 1:2 and DNA to Booster ratio of 1:3. Transfection complexes were removed 16 hours post transfection and replaced with fresh growth media. Ten days post transfection the cells were harvested by scraping into the media and subjected to three rounds of freeze-thaw using LN2. For protocol one, crude virus stocks were used to infect 8 x T175 of 70% confluent Vero cells at 37 °C, after 2 hours the virus was removed and replaced with fresh growth media. For protocol two, crude virus stocks were used to infect 36 x T175 flasks of 50% confluent Vero cells at 37°C; after 2 hr the virus was removed and replaced with fresh growth media.

2.3.4 BKV purification

2.3.4.1 Protocol one

Virions were harvested 14 days post infection and purified as described previously (M. Jiang et al., 2009). Briefly, Vero cells were harvested by scraping and pelleted by centrifugation at 8000 g, the pellet was resuspended in buffer A1 (10 mM HEPES pH 7.9, 1 mM CaCl2, 1 mM MgCl2, 5 mM KCl), sonicated in a water bath for 5 min and treated with 1 U/ml neuraminidase (Sigma) for 1 hour at room temperature (RT), after having the pH adjusted to 6.0. The pH was adjusted back to 7.4 and the lysate heated at 40 °C for 5 min. The lysate was pelleted at 16,000 g for 5 min and then the pellet resuspended in buffer A1 and incubated with 0.1% (w/v) deoxycholic acid for 15 min before being pelleted again. The supernatants were combined over a 4 ml 20% (w/v) sucrose cushion in buffer A1 before centrifugation at 85,000 g for 3 hours at 4 °C in a Beckman SW32Ti rotor. The pellet was resuspended in 1 ml buffer A1 and layered over a 6 ml preformed 1.2-1.4 g.cm⁻³ CsCl gradient before centrifugation at 155,000 g for 16 hours at 15 °C in a Beckman SW40Ti rotor. The band containing mature virions was collected using a 26-guage needle and dialysed against buffer A1 overnight at 4 °C. For structural analysis by cryo-EM virions were concentrated 20-fold in a Centrifugal Concentrator with a 1 MDa cut off at 4 °C for 15 mins (Vivaspin® 500, Vivaproducts).
2.3.4.2 Protocol two

BKV-infected Vero cells were harvested 21 days post infection and purified essentially as described previously (Shen et al., 2011). Cells were harvested by scraping and pelleted at 4000 g before being resuspended in 10 ml of buffer A2 (10 mM Tris, 50 mM NaCl, 0.01% Triton X-100) supplemented with an EDTA-free protease inhibitor cocktail (Roche). The lysate was freeze thawed three times with 3 min sonication in a water bath between each cycle. Deoxycholic acid was added to a final concentration of 0.25% and incubated for 30 minutes at 37 °C with shaking. The pH was then lowered to 6.0 with 0.5 M HEPES (pH 5.4) and 5 units of type V neuraminidase (Sigma) was added. This was incubated at 37 °C for 1 hour with shaking and the pH was raised back to 7.5 with 0.5 M HEPES (pH 8.0). The sample was then sonicated for 3 x 45 seconds in a water bath before pelleting the cellular debris at 4000 g for 10 minutes. The pellet was resuspended in 5 ml of buffer A2 and this process was repeated a further two times. The supernatants were combined over a 4 ml 20% (w/v) sucrose cushion in buffer A2 before centrifugation at 85,000 g for 3 hours at 4 °C in a SW32Ti rotor (Beckman). The pellet was resuspended in 5 ml of 1.34 g/ml CsCl and the isopycnic gradient was spun at 4 °C for 16 hours at 110,000 g (no brake) in a SW55Ti rotor (Beckman). The BKV band was collected using a 26-gauge needle and dialysed against 2 L of buffer A2 (without Triton X-100) for two days at 4 °C. Dialysis buffer was exchanged twice with pre-chilled buffer A2 (without Triton X-100). The concentration of purified BKV was determined using a NanoDrop spectrophotometer (Thermo Fisher Scientific). Buffer A2 (w/o Triton X-100) was used to take a blank reading and then a 1 μl aliquot of virus solution was applied to the sample pedestal to take an $A_{280}$ reading. The quality of the purified sample was assessed by negative stain electron microscopy analysis (see 2.4.1).

2.3.5 Virus-like particle expression and purification

Virus-like particles (VLPs) were prepared by Ethan Morgan, University of Leeds, using an updated and slightly modified protocol that has been previously described (Touzé et al., 2001; Buck et al., 2004). Briefly, 293TT cells were co-transfected with a codon modified BKV expression plasmids for VP1 (plaw), together with a reporter plasmid encoding eGFP, (kind gift from Dr Chris Buck, National Cancer Institute, Bethesda, USA). 48 h after transfection, VLPs were harvested and re-suspended at a concentration of >100 million/ml in buffer A1.
Neuraminidase V (Sigma) was added to a final concentration of 1 U/ml, and the cell suspension was incubated for 15 min at 37 °C, followed by cell lysis by the addition of 0.5% Triton X-100 (Sigma) and further incubation at 37 °C for an additional 15 min. The lysate was then treated with 0.1% benzonase (Sigma) and 0.1% DNase (Sigma). Capsid maturation was allowed to proceed overnight at 37 °C. Lysates were clarified by centrifugation at 5,000 g for 10 min and purified through a 27 to 33 to 39% iodixanol gradient (Optiprep; Sigma). After purification, fractions were collected by gravity flow and a 10 μL sample was analysed for VP1 expression by Western blot. Protein concentrations for the VLPs were calculated using a bicinchoninic acid (BCA) assay according to the manufacturer’s instructions (Pierce, USA).

To determine the infectivity of the BKV VLPs, HEK293TT cells were seeded at 150,000 cells per well in 6-well plates, 24 hours prior to infection. 5 μg of purified BKV VLPs (approximately 100-fold more than reported previously (Schowalter and Buck, 2013)) was added to the wells and incubated overnight. Cells were then analysed by microscopy for eGFP expression. For structural analysis by cryo-EM, VLPs were concentrated 20-fold in a Centrifugal Concentrator with a 1 MDa cut off at 4 °C for 15 mins (Vivaspin® 500, Vivaproducts).

2.3.6 Immunofluorescence

Immunofluorescence experiments were performed by Dr Emma Prescott and Margarita Panou, University of Leeds. RPTE cells were grown on glass coverslips and infected at around 50% confluence with 100 genomes/cell of purified virus in opti-MEM. The cells were fixed three days post infection with 4% paraformaldehyde in PBS for 10 min, followed by permeabilisation with 0.1% Triton in PBS for 5 min. The cells were stained with mouse anti-VP1 P5G6 (a gift from Professor Denise Galloway, Fred Hutchinson Cancer Research Centre) used 1:500 and rabbit anti-SV40 VP2 + VP3 (Abcam) used 1:250, followed by Alexa Fluor® 488 chicken anti-mouse and Alexa Fluor® 594 chicken anti-rabbit (Life Technologies) used 1:500. ProLong® Gold antifade mountant with DAPI (Life Technologies) was used to mount the coverslips, which were imaged using a Zeiss LSM700 inverted confocal microscope.
2.3.7 DNA extraction from VLPs and Virions

For analysis of packaged DNA in VLPs and virions, 50 μL of purified VLPs or virions were incubated at 56 °C in water containing 5 μL each of Proteinase K, 10% SDS and 0.5 M EDTA. After incubation, half of the sample was digested with 1 μL BamHI at 37 °C for 2 hours to allow analysis of packaged genome or plasmid DNA and cellular DNA. 5 μL of digested sample was then added to gel loading dye and loaded onto a 0.7% agarose gel for DNA analysis – see Agarose gel electrophoresis (2.3.8). As a control for cellular DNA, 100 ng of DNA extracted from HEK293TT cells using E.Z.N.A.® Tissue DNA Kit (Omega Bio-Tek) was loaded on the gel.

2.3.8 Agarose gel electrophoresis

Individual DNA samples (5 μl) were combined with 10 μl of 5 x DNA loading buffer and loaded into wells on a 0.7% agarose gel. HyperLadderTM 1 kb - Bioline (6 μl) was also loaded into one of wells to allow for easy size determination of DNA bands. The samples were run in a Mini-Sub® Cell GT (Bio-Rad). These were run at 80 Volts for 60 minutes in Tris-acetate-EDTA (TAE) buffer. A 10 μl aliquot of stock SYBR® Safe DNA Gel Stain (Invitrogen™) was added to the TEA buffer.

2.3.9 Fluorescent Focus Unit Assay

RPTE cells were seeded out into 96-well-plate (2 × 103 cells per well, in a total volume of 100 μL) and incubated for 16 h. Purified BKV was serially diluted two-fold into serum-free media (in a total volume of 100 μL per well) and allowed to infect RPTE cells for 2 h at 37 °C. Infected cells were washed once with PBS and fresh media was added. RPTE cells were incubated for 48 h at 37 °C. Cells were fixed with 4% paraformaldehyde for 10 min at RT and washed with PBS. Fixed RPTE cells were permeabilised with 0.1% Triton-X100 (Sigma-Aldrich, St-Louis, MI, USA) in PBS, washed and incubated overnight at 4 °C in primary antibody against VP1 protein. Anti-VP1 primary antibody was used at 1:250 dilutions (in PBS with 1% BSA). Cells were further washed and incubated with a fluorophore-488-conjugated chicken anti-mouse secondary antibody (1:250 in PBS with 1% BSA) for 1 h at 37 °C. Finally, RPTE cells stored in PBS and the plate was imaged with the IncuCyte ZOOM instrument (Essen BioScience, Ann Arbor, MI, USA). The software parameters with a 10× objective was used for imaging (Stewart et al., 2015). IncuCyte analysis was performed with assistance from Margarita.
Panou, University of Leeds. The number of positive infected cells per well was calculated. BKV titre was measured by multiplying the number of positive-infected cells/well by the corresponding dilution factor and the IU mL$^{-1}$ was determined by calculating the number of infected cells in the entire well from the mean number of infected cells in the 10 fields of view, and then the number of infectious units calculated (Stewart et al., 2015).

2.3.10 Western blotting

Triton lysis buffer (10mM Tris pH 7.6, 10mM sodium phosphate, 130 mM NaCl, 1% Triton X-100, 20 mM N-Ethylmaleimide, complete protease inhibitor cocktail (Roche)) was used to harvest total cellular protein from BKV infected cells. Proteins from cell lysates or purified virus were separated on 10% SDS PAGE gels, with 20 µg of protein loaded per sample. Proteins were then transferred by a semi-dry transfer method (Trans Blot® SD Semi-Dry Transfer cell, Bio-Rad, USA) onto a nitrocellulose membrane (GE Healthcare, USA). Membranes were blocked with 5% milk solution and the following primary antibodies were used at a dilution of 1:5000: VP1 P5G6 (a gift from Professor Denise Galloway, Fred Hutchinson Cancer Research Centre); mouse anti-VP1 p5G6 used 1:5000, rabbit anti-SV40 VP2 + VP3 (Abcam; ab53983) used 1:1000 and mouse anti-GAPDH (Santa Cruz) used 1:5000. HRP-conjugated mouse or rabbit secondary antibodies (Sigma, USA) were then used at a 1:5000 dilution. Proteins were detected using WesternBright ECL (Advansta, USA) and visualised on X-ray film.

2.3.11 Quantitative PCR

Quantitative PCR (qPCR) measurements were performed with assistance from Dr Emma Prescott, University of Leeds. Total DNA was extracted from the infected cells using the E.Z.N.A.® Tissue DNA kit (Omega Bio-Tek) and 10 ng of DNA was analysed by qPCR using the QuantiFast SYBR Green PCR kit (Qiagen) with the following primers against BKV Dunlop; BKV Dunlop copy Forward TGTGATTGGGATTCAGTGCT and BKV Dunlop copy Reverse AAGGAAAGGCTGGATTCTGA. A serial dilution of the pGEM7 Dunlop plasmid was used to calculate the copy number per µg of total DNA.
2.4 Structural methods

2.4.1 Negative Stain Electron Microscopy

For all negative staining, 3 µL aliquots of BKV or VLP were applied to 300 mesh copper formvar/carbon grids (Agar Scientific, UK) that had been glow-discharged for 30 s in air. Samples were then stained with 1% uranyl acetate solution, and allowed to dry in air for ~5 mins. Samples were imaged on a Tecnai G²-Spirit transmission EM at 120 keV, equipped with a Gatan US1000XP CCD camera.

2.4.2 Cryo-EM sample preparation and data acquisition

2.4.2.1 Intermediate resolution BKV studies

For cryo-EM, 3 µL aliquots of BKV or VLP in buffer A1 were applied to Quantifoil R2/1 200 mesh Cu grids EM grids that had been glow discharged in air for ~30 s. The samples were then blotted in 100% relative humidity and vitrified by plunging into liquid nitrogen-cooled liquid ethane, using a FEI Vitrobot mark IV or Leica-EM-GP. Wild-type BKV was imaged using a FEI Tecnai-F20 transmission electron microscope at 200 keV with a Gatan 626 holder, and a Gatan K2-Summit direct electron detector. Images were recorded under low-dose conditions using the K2 camera in counting mode. Images were recorded at 19,000x magnification, at a dose rate of ~4 e²/Å²/s, a 4 frames per second frame rate, and a 10 s exposure. The recorded images had a calibrated object sampling of 1.92 Å/pixel and a total accumulated dose of ~40 e²/Å². For the VLP sample, grids were imaged at OPIC on a FEI Polara Microscope. Images were obtained on a Gatan K2 summit camera operating in super resolution mode, with a final super resolution pixel sampling of 0.675 Å/pix and a physical pixel sampling of 1.35 Å/pix.

2.4.2.2 High-resolution BKV studies

Cryo-EM grids were prepared by applying 3 µl of purified BKV (0.25 mg/ml) in buffer A2 to 400 mesh lacey grids coated in a 3 nm carbon film (Agar Scientific, UK). The sample was left to adsorb for 30 seconds before most of the sample was blotted away manually and this process was repeated four times. For the reduced BKV structure, 5 mM DTT was added to the sample and incubated for 15 minutes at RT prior to preparing cryo-grids. On-grid binding of receptors was performed by applying 3 µl of a 20 mM solution of GT1b oligosaccharide
(CarboSynth) or heparin 10 mg/ml (Sigma – 4784) in Buffer A2 (without Triton X-100) to the pre-blotted, BKV-coated grid, and leaving for 30 seconds before blotting and freezing using a Leica EM GP plunge freeze device (Leica Microsystems). The Leica EM chamber temperature was set to 8 °C with 80% relative humidity and liquid ethane was used for sample vitrification. Grids were glow discharged for 30 seconds prior to application of the samples. All data sets were collected on a FEI Titan Krios (ABSL, University of Leeds) transmission electron microscope at 300 kV, at a magnification of 75,000x and a final calibrated object sampling of 1.065 Å/pixel. Exposures were recorded using the EPU automated acquisition software on a FEI Falcon II (unliganded BKV) or Falcon III direct electron detector operating in linear mode. Detailed information on data collection is shown in Table 2-3.

2.4.2.3 CV-A24v studies

The CV-A24 sample was provided by Dr Georg Zocher, University of Tübingen. Cryo-EM grids were prepared by applying 3 µl of purified CV-A24v (10 mg/ml) to 400 mesh lacy grids coated in a 3 nm carbon film (Agar Scientific, UK). The sample was left to adsorb for 30 seconds before most of the sample was blotted away manually. On-grid binding of the receptor was performed by applying 3 µl of ICAM-1 D1-D2 (9.85 mg/ml) to the pre-blotted, virus-containing grid, and leaving for 30 seconds before blotting and freezing using a Leica EM GP plunge freeze (Leica Microsystems) device. The Leica EM chamber temperature was set to 8 °C with 80% relative humidity and liquid ethane was used for sample vitrification. Grids were glow discharged for 30 seconds prior to application of the samples. All CV-A24v data sets were collected on FEI Titan Krios (ABSL, University of Leeds) transmission electron microscopes at 300 kV. Exposures were recorded using the EPU automated acquisition software on a FEI Falcon III direct electron detector operating in linear mode or on a Gatan K2 summit camera operating in super resolution mode. Detailed information on data collection is shown in Table 2-4.

2.4.3 Image processing

2.4.3.1 Intermediate resolution BKV studies

For BKV virions, a total of 432 micrographs were recorded, each consisting of a movie of 40 frames. Images were converted from Gatan DM4 format to MRC
using the dm2mrc tool in IMOD (Kremer et al., 1996). A drift correction calculation was then performed on each stack of movie frames using Motioncor (X. Li et al., 2013), writing out a drift-corrected average of frames 5-25. This was done to eliminate both the earliest frames (where beam-induced movement is strongest), and the last frames (where radiation damage predominates). The defocus of each drift-corrected average was then determined using Cttffind3 (Mindell and Grigorieff, 2003), and were in the range 0.6-4.8 microns. All further image processing steps were performed in Relion1.3 (Scheres, 2012). Particles were picked manually and yielded a total dataset of 7697 virions. Using iterative 3D classification into two classes, a subset of 2237 images was taken forward for 3D reconstruction. The resulting map was sharpened using a B-factor of -456 Å² to give a final resolution of 7.6 Å using a ‘gold-standard’ Fourier Shell Correlation between two independently refined halves of the dataset at 0.143.

For BKV VLPs, a total of 170 micrographs were recorded, each consisting of a movie of 40 frames. A drift correction calculation was then performed on each stack of movie frames using Motioncor (X. Li et al., 2013). The defocus of each drift-corrected micrograph was determined using Cttffind3 (Mindell and Grigorieff, 2003). All further image-processing steps were performed in Relion1.3 (Scheres, 2012). Particles were interactively selected from all micrographs and yielded a total dataset of 2,888 particles. Using iterative 3D classification into two classes, a subset of 2517 images was taken forward for 3D reconstruction. The resulting map was sharpened using a B-factor of -804 Å² and using the ‘gold-standard’ Fourier Shell Correlation between two independently refined halves of the dataset at 0.143, the final resolution was 9.1 Å.

2.4.3.2 High-resolution BKV and CV-A24v studies

Image processing was carried out using the Relion2.0 pipeline (Kimanius et al., 2016). Drift-corrected averages of each movie were created using Motioncor2 (Zheng et al., 2017) and the contrast transfer function of each determined using gctf (K. Zhang, 2016). Automated particle picking on lacey carbon grids resulted in a large number of boxes picked on the edges of holes in the carbon film. To remove such ‘junk’ particles from the data sets, 2D classification in Relion2.0 was used with CTF amplitude correction only performed from the first peak of each CTF onwards. Particles were further classified using several rounds of both reference-free 2D classification and 3D classification. After each round, the best
classes/class was taken to the next step of classification. Icosahedral symmetry was imposed during 3D auto-refinement and post-processing was employed to appropriately mask the model, estimate and correct for the B-factor of the maps. The final resolutions were determined using the ‘gold standard’ Fourier shell correlation criterion (FSC = 0.143) shown in Table 2-3. Local resolution was estimated in Relion2.0 which also generated maps filtered by local resolution.
Table 2-3 – Data collection and imaging processing parameters for BKV structures.

<table>
<thead>
<tr>
<th>Microscope</th>
<th>BKV</th>
<th>BKV-VLP</th>
<th>BKV</th>
<th>BKV-GT1b</th>
<th>BKV-Heparin</th>
<th>BKV-DTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera</td>
<td>FEI Tecnai F20</td>
<td>FEI Polara</td>
<td>FEI Titan Krios</td>
<td>FEI Titan Krios</td>
<td>FEI Titan Krios</td>
<td>FEI Titan Krios</td>
</tr>
<tr>
<td>Voltage</td>
<td>200 keV</td>
<td>300 keV</td>
<td>300 keV</td>
<td>300 keV</td>
<td>300 keV</td>
<td>300 keV</td>
</tr>
<tr>
<td>Pixel Size</td>
<td>1.92</td>
<td>1.35</td>
<td>1.065</td>
<td>1.065</td>
<td>1.065</td>
<td>1.065</td>
</tr>
<tr>
<td>Total Dose</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>59</td>
<td>55</td>
<td>66</td>
</tr>
<tr>
<td>Number of Frames</td>
<td>40</td>
<td>40</td>
<td>39</td>
<td>59</td>
<td>39</td>
<td>45</td>
</tr>
<tr>
<td>Defocus Range</td>
<td>-0.5 – 3.5</td>
<td>-0.5 – 3.5</td>
<td>-0.5 - 3.0</td>
<td>-0.5 - 3.0</td>
<td>-0.5 - 2.3</td>
<td>-0.5 – 3.0</td>
</tr>
<tr>
<td>Micrographs</td>
<td>432</td>
<td>170</td>
<td>3992</td>
<td>8651</td>
<td>6227</td>
<td>3474</td>
</tr>
<tr>
<td>Acquisition software</td>
<td>Gatan DM</td>
<td>Gatan DM</td>
<td>FEI EPU</td>
<td>FEI EPU</td>
<td>FEI EPU</td>
<td>FEI EPU</td>
</tr>
<tr>
<td>Motion correction</td>
<td>MotionCor</td>
<td>MotionCor</td>
<td>MotionCor2</td>
<td>MotionCor2</td>
<td>MotionCor2</td>
<td>MotionCor2</td>
</tr>
<tr>
<td>CTF estimation</td>
<td>Ctffind3</td>
<td>Ctffind3</td>
<td>Gctf</td>
<td>Gctf</td>
<td>Gctf</td>
<td>Gctf</td>
</tr>
<tr>
<td>Software</td>
<td>Relion1.3</td>
<td>Relion1.3</td>
<td>Relion2.0</td>
<td>Relion2.0</td>
<td>Relion2.0</td>
<td>Relion2.0</td>
</tr>
<tr>
<td>Particles contributed</td>
<td>2,237</td>
<td>2,517</td>
<td>11,980</td>
<td>40,334</td>
<td>24,791</td>
<td>8,733</td>
</tr>
<tr>
<td>B-factor</td>
<td>-456</td>
<td>-804</td>
<td>-174</td>
<td>-212</td>
<td>-206</td>
<td>-326</td>
</tr>
<tr>
<td>Resolution (FSC 0.143)</td>
<td>7.6</td>
<td>9.1</td>
<td>3.8</td>
<td>3.4</td>
<td>3.6</td>
<td>6.5</td>
</tr>
</tbody>
</table>
Table 2-4 - Data collection and imaging processing parameters for CV-A24v structures.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Microscope</td>
<td>FEI Titan Krios</td>
<td>FEI Titan Krios</td>
<td>FEI Titan Krios</td>
<td>FEI Titan Krios</td>
<td>FEI Titan Krios</td>
<td>FEI Titan Krios</td>
</tr>
<tr>
<td>Camera</td>
<td>Falcon III</td>
<td>Gatan K2 Summit</td>
<td>Gatan K2 Summit</td>
<td>Falcon III</td>
<td>Falcon III</td>
<td>Falcon III</td>
</tr>
<tr>
<td>Voltage</td>
<td>200 keV</td>
<td>300 keV</td>
<td>300 keV</td>
<td>300 keV</td>
<td>300 keV</td>
<td>300 keV</td>
</tr>
<tr>
<td>Pixel Size</td>
<td>1.065</td>
<td>0.82</td>
<td>0.82</td>
<td>1.065</td>
<td>1.065</td>
<td>1.065</td>
</tr>
<tr>
<td>Total Dose</td>
<td>60</td>
<td>41</td>
<td>41</td>
<td>63</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>Number of Frames</td>
<td>40</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>59</td>
<td>59</td>
</tr>
<tr>
<td>Defocus Range</td>
<td>-0.5 – 3.0</td>
<td>-0.5 – 3.0</td>
<td>-0.5 – 3.0</td>
<td>-0.5 – 2.5</td>
<td>-0.5 – 2.5</td>
<td>-0.5 – 2.5</td>
</tr>
<tr>
<td>Micrographs</td>
<td>2652</td>
<td>1514</td>
<td>1514</td>
<td>11479</td>
<td>11479</td>
<td>11479</td>
</tr>
<tr>
<td>Acquisition software</td>
<td>FEI EPU</td>
<td>FEI EPU</td>
<td>FEI EPU</td>
<td>FEI EPU</td>
<td>FEI EPU</td>
<td>FEI EPU</td>
</tr>
<tr>
<td>Motion correction</td>
<td>MotionCor2</td>
<td>MotionCor2</td>
<td>MotionCor2</td>
<td>MotionCor2</td>
<td>MotionCor2</td>
<td>MotionCor2</td>
</tr>
<tr>
<td>CTF estimation</td>
<td>Gctf</td>
<td>Gctf</td>
<td>Gctf</td>
<td>Gctf</td>
<td>Gctf</td>
<td>Gctf</td>
</tr>
<tr>
<td>Software</td>
<td>Relion2.0</td>
<td>Relion2.0</td>
<td>Relion2.0</td>
<td>Relion2.0</td>
<td>Relion2.0</td>
<td>Relion2.0</td>
</tr>
<tr>
<td>Particles contributed</td>
<td>26,311</td>
<td>31,381</td>
<td>1,206</td>
<td>115,601</td>
<td>1,049</td>
<td>1,218</td>
</tr>
<tr>
<td>B-factor</td>
<td>-257</td>
<td>-133</td>
<td>-120</td>
<td>-212</td>
<td>-236</td>
<td>-223</td>
</tr>
<tr>
<td>Resolution (FSC 0.143)</td>
<td>3.9</td>
<td>3.0</td>
<td>3.8</td>
<td>3.2</td>
<td>4.3</td>
<td>4.3</td>
</tr>
</tbody>
</table>
2.4.4 Modelling and refinement

2.4.4.1 BKV homology modelling and flexible fitting

A homology model of the BKV VP1 asymmetric unit based on the crystal structure of SV40 (PDB 1SVA) was built using the SWISS-MODEL server. This was then fitted (as a rigid body) into a corresponding segment of the virion cryo-EM density map generated using UCSF Chimera (Pettersen et al., 2004). Flexible fitting of the homology model was then carried out using MDFF (Trabuco et al., 2011).

2.4.4.2 BKV model building and refinement

A preliminary model for the BKV asymmetric unit was generated by rigid body fitting the previously published homology model (5FUJA) (Hurdiss et al., 2016) into the EM density map using UCSF Chimera (Pettersen et al., 2004). The model was then manually fitted in Coot using the ‘real space refinement tool’ (Emsley and Cowtan, 2004). The resulting model was then symmetrised in Chimera to generate the capsid and subject to refinement in Phenix (Headd et al., 2012). Iterative rounds of manual fitting in Coot and refinement in Phenix were carried out to improve non-ideal rotamers, bond angles and Ramachandran outliers. For VP2, the murine polyomavirus (PyV) coordinates (1CN3) were aligned to a five-fold coordinated penton of BKV using UCSF Chimera. The VP2 model was then mutated, fitted and refined in Coot before being subject to refinement in Phenix (Table 2-5). For GT1b modelling, the coordinates for the oligosaccharide component from the previously determined BKV penton – GD3 structure (Neu et al., 2013) were used to model the disialic acid motif from the ‘right’ arm of GT1b. Low-pass filtered maps were generated using Relion image handler and difference maps were generated in UCSF Chimera. Interactions between chains 6 and 3 were analysed using PDBBePISA (Krissinel and Henrick, 2007).
Table 2-5 – Quality assessment of the BKV-GT1b atomic model.

<table>
<thead>
<tr>
<th>BKV-GT1b model statistics</th>
<th>Phenix real space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refinement software</td>
<td></td>
</tr>
<tr>
<td>Map CC (whole unit cell)</td>
<td>0.81</td>
</tr>
<tr>
<td>Map CC (all atoms)</td>
<td>0.83</td>
</tr>
<tr>
<td>RMSD bond lengths</td>
<td>0.01</td>
</tr>
<tr>
<td>RMSD bond angles</td>
<td>0.93</td>
</tr>
<tr>
<td>Ramachandran preferred</td>
<td>91.93%</td>
</tr>
<tr>
<td>Ramachandran allowed</td>
<td>8.07%</td>
</tr>
<tr>
<td>Ramachandran outlier</td>
<td>0.00%</td>
</tr>
<tr>
<td>Rotamer outliers</td>
<td>0.22%</td>
</tr>
<tr>
<td>C-beta deviations</td>
<td>0</td>
</tr>
<tr>
<td>All-atom clash score</td>
<td>3.36</td>
</tr>
</tbody>
</table>

2.4.4.3 CV-A24v model building and refinement

A preliminary model was generated by rigid body fitting the asymmetric unit from the X-ray structure of CV-A24v (4Q4W (Zocher et al., 2014)) and for ICAM-1, the D1 domain from (1Z7Z (Xiao et al., 2005)) into the EM density map using UCSF Chimera (Pettersen et al., 2004). The model was then manually fitted in coot using the ‘real space refinement tool’ (Emsley and Cowtan, 2004). The resulting model was then symmetrised in Chimera to generate the capsid and subject to refinement in Phenix (Headd et al., 2012). Iterative rounds of manual fitting in coot and refinement in Phenix were carried out to improve non-ideal rotamers, bond angles and Ramachandran outliers (Table 2-6). The asymmetric unit was assessed for quality using MolProbity (V.B. Chen et al., 2010). Interactions between ICAM-1 D1 and VP1-3 were analysed using PDBePISA (Krissinel and Henrick, 2007) and the roadmap projection of the CV-A24v surface was generated using RIVEM (Xiao and Rossmann, 2007).
Table 2-6 - Quality assessment of CV-A24v atomic models.

<table>
<thead>
<tr>
<th>CV-A24v model statistics</th>
<th>CV-A24v-ICAM-1</th>
<th>CV-A24v</th>
<th>CV-A24v Empty particle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refinement software</td>
<td>Phenix real space</td>
<td>Phenix real space</td>
<td>Phenix real space</td>
</tr>
<tr>
<td>Map CC (whole unit cell)</td>
<td>0.75</td>
<td>0.76</td>
<td>0.83</td>
</tr>
<tr>
<td>Map CC (all atoms)</td>
<td>0.87</td>
<td>0.85</td>
<td>0.82</td>
</tr>
<tr>
<td>RMSD bond lengths</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>RMSD bond angles</td>
<td>0.87</td>
<td>0.91</td>
<td>1.04</td>
</tr>
<tr>
<td>Ramachandran preferred</td>
<td>94.09%</td>
<td>95.75%</td>
<td>92.25%</td>
</tr>
<tr>
<td>Ramachandran allowed</td>
<td>5.91%</td>
<td>4.13%</td>
<td>7.57%</td>
</tr>
<tr>
<td>Ramachandran outlier</td>
<td>0.00%</td>
<td>0.12%</td>
<td>0.18%</td>
</tr>
<tr>
<td>Rotamer outliers</td>
<td>0.40%</td>
<td>0.00%</td>
<td>0.40%</td>
</tr>
<tr>
<td>C-beta deviations</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>All-atom clash score</td>
<td>4.37</td>
<td>3.15</td>
<td>4.63</td>
</tr>
</tbody>
</table>

2.5 Molecular dynamics simulations

Molecular dynamic (MD) simulations were performed by Dr Martin Frank, Biognos AB, Sweden. The X-ray crystal structure 4MJ0 (Neu et al., 2013), was used as a starting structure for MD. Missing protein loops were modelled with YASARA (Krieger and Vriend, 2015). The GT1b oligosaccharide was built with AMBER tleap using Glycam06 building blocks (Kirschner et al., 2008). The glycosidic torsions of the (2-8)-linkage were adjusted to the values found in the x-ray structure for GD3 in site C. GT1b molecules were positioned into the four occupied binding sites by superimposing the terminal Neu5Ac (R) with the terminal Neu5Acs present in the X-ray structure, respectively. An additional 5th GT1b molecule was manually positioned into the remaining binding site resembling the same binding mode. During very long MD simulations, dissociation of ligands can occur, particularly when the binding mode is not optimised at the beginning of the simulation as is the case for 'docked ligands'.
Since the binding mode of the terminal Neu5Ac (R) is well established by several H-bonds, distance restraints (Asn-272:OD1-Neu5Ac:N5, Ser-274:OG-Neu5Ac:O1A and Thr-276:OG1-Neu5Ac:O1B) were applied in order to guarantee that this residue stayed in the binding site in a defined binding mode for the whole simulation. A periodic box was filled with a 0.1% NaCl solution, which resulted in a molecular system consisting of 126430 atoms. The simulations were performed at 310 K with YASARA in ‘fast mode’ using AMBER14 (which includes Glycam06) as a force field. In total 1.1 µs trajectory data was sampled for the BKV-GT1b pentamer complex distributed over six independent simulations. This allowed analysis of GT1b dynamics in the binding site(s) of BKV over an accumulated time of 5.5 µs. Snapshots were recorded every 50 ps, which gave a trajectory file consisting of 22200 frames and produced about 20 GByte binary trajectory data in XTC format. All analysis and generation of scientific plots was performed with Conformational Analysis Tools (CAT) - www.md-simulations.de/CAT/. For calculation of the atom 3D density map all MD frames of the pentamer were superimposed on the EM atomic model using a sequence alignment file. In order to average the density over the five sides the five-fold symmetry was considered. Snapshots were visualised with VMD (Humphrey et al., 1996).

2.6 Visualisation

All figures were generated using UCSF Chimera and PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).
Chapter 3 Structural studies of BK polyomavirus

3.1 Introduction

Polyomaviruses are small, non-enveloped, dsDNA viruses which use mammals, birds and fish as their natural hosts (White et al., 2013; Buck et al., 2016). The first two human polyomaviruses discovered, BKV and JC polyomavirus (JCV), were named after the index case-patients upon their discovery more than 40 years ago (Padgett et al., 1971; Gardner et al., 1971). The last decade has seen the discovery of many new polyomaviruses linked to human disease, including Trichodysplasia spinulosa-associated polyomavirus (TSV) (van der Meijden et al., 2010) and Merkel Cell polyomavirus (MCV) (H. Feng et al., 2008) which cause skin lesions and an aggressive skin carcinoma respectively. These discoveries have led to a resurgence of interest in polyomavirus biology (DeCaprio and Garcea, 2013).

Following a recent revision in the taxonomy of polyomaviruses, BKV now belongs to the Betapolyomavirus genus (Figure 3-1) within the Polyomaviridae family (Moens et al., 2017). There are four major genotypes of BKV (I, II, III and IV), and these are further divided into subgroups (Jin et al., 1993). BKV genotype I has a worldwide distribution, genotype II and III are infrequently observed, and genotype IV is found primarily in Europe and East Asia (Yogo et al., 2009). Serological studies have demonstrated that each of the BKV genotypes and subgroups Ib1 and Ib2 represent distinct serotypes (Pastrana et al., 2013).

BKV is an opportunistic pathogen, capable of causing several diseases in the immunosuppressed (Knowles, 2006). Infection with BKV typically occurs in childhood, may occur through a respiratory or oral route, and is either asymptomatic or causes a mild respiratory illness (Goudsmit et al., 1982). A persistent infection is then established in the kidney and urinary tract, and 80-90% of the general population are thought to be seropositive for the most prevalent BKV genotypes (Antonsson et al., 2010; Egli et al., 2009; Knowles et al., 2003; Stolt et al., 2003). However, in the immunosuppressed, BKV can become ‘reactivated’, resulting in shedding into the urine because of increased replication in the absence of competent immune surveillance and control (Ahsan and Shah, 2006). Such an increase in replication is ultimately linked with severe health problems, including polyomavirus-associated nephropathy (PVAN) (Balba
et al., 2013) and haemorrhagic cystitis (HC) (Dropulic and Jones, 2008) in patients who have received kidney and bone marrow transplants, respectively. For example, up to 10% of kidney transplant patients experience PVAN, and up to 90% of these will go on to lose their graft (Ramos et al., 2009). Generic antiviral drugs such as Cidofovir have been used to treat BKV-associated diseases, but have low efficacy and are themselves associated with nephrotoxicity (Safrin et al., 1997). Treatment is typically limited to a reduction in immunosuppression, which runs substantial risks of transplant rejection (Kuypers, 2012; Vats et al., 2006).

The incidence of BKV-related disease is rising owing to the increasing number of transplants, and the immunosuppressive drug regimens used to support such patients (Bennett et al., 2012). In addition, there is also mounting evidence that BKV, like MCV, may be carcinogenic in humans (D.C. Müller et al., 2018; Sandy Liu et al., 2017; Papadimitriou et al., 2016; Luo et al., 2017; Frascà et al., 2015; H. Feng et al., 2008). No effective antivirals specifically targeting BKV are available (Bennett et al., 2012; Ambalathingal et al., 2017), and so there is increasing interest in the development of targeted therapeutics. Such therapies may include vaccination with a VLP vaccine, intravenous immunoglobulin therapy or antiviral compounds designed to interfere with capsid assembly or cellular entry. High-resolution structural information for the infectious BKV virion and its interaction with cellular receptors would be a useful resource in designing and evaluating such therapies.
Figure 3-1 – Polyomaviridae. Phylogenetic relationships of polyomaviruses based on conserved amino acid blocks of the TAg coding sequence. Polyomaviruses are denoted by species names and Genbank accession numbers, and their genome sizes are given on the right-hand side of the figure. They are grouped into genera by colouring. (Figure taken from (Moens et al., 2017) [published under creative commons {CC BY-SA 4.0} license].)
3.1.1 Genome organisation and replication

The genome of BKV is a closed, circular dsDNA molecule of approximately 5,300 bp in length (Figure 3-2). It comprises early and late coding regions which are separated by a non-coding control region (NCCR) which contains a unique origin from which bidirectional genome replication can occur. Significant variation in the NCCR is observed between different BKV isolates. “Archetypal strains” are believed to be the transmissible form of the virus which circulate in human populations and are commonly found in urine (R. Rubinstein et al., 1987; Yogo et al., 2008). Isolates with a rearranged NCCR are usually found in kidneys or other disease tissue (Cubitt, 2006; Helle et al., 2017; Ambalathingal et al., 2017). The early region of the BKV genome encodes small T antigen (tAg), truncated T antigen (truncTAg) and large T antigen (TAg) which are each involved in viral replication. Transcription of the early proteins occurs before replication resulting in the production of tAg & TAg which accumulate in the nucleus to promote replication of the viral genome (Carbone et al., 1997; Sriaroon, 2010). TAg forms a multimeric AAA+ ATPase with helicase activity which binds to the origin of replication to promote transcription of the late coding regions (W. Wang and D.T. Simmons, 2009; Mastrangelo et al., 1989). TAg is also able to bind to tumour suppressor proteins p53, p130, p107 and Rb and by doing so can drive the host cell cycle into the S phase (K.F. Harris et al., 1998; K.F. Harris et al., 1996; Pipas, 1992; Shivakumar and G.C. Das, 1996). The tAg is involved in transformation, cell cycle progression and viral replication (Cicala et al., 1994; Howe et al., 1998; Whalen et al., 1999). After replication of the viral genome, the late region of the viral genome is expressed to produce the structural proteins VP1, VP2 and VP3 and the agnoprotein (Cubitt, 2006). Following production in the cytoplasm, VP1-3 localise to the nucleus where virion assembly occurs (Drachenberg et al., 2003; T.-C. Li et al., 2003; Bennett et al., 2015). The agnoprotein is a small, highly basic protein which has recently been shown to play a role in BKV egress (Panou et al., 2018).
3.1.2 Virion structure

3.1.2.1 Capsid

Much of our current understanding of polyomavirus structure comes from the pioneering work of Caspar (J.P. Griffith et al., 1992; Rayment et al., 1982) and Harrison (Stehle et al., 1996; Stehle and Harrison, 1996) on the archetypal polyomaviruses SV40, and PyV. The crystal structures of these viruses revealed that polyomavirus capsids consist of 360 copies of the major capsid protein VP1. These VP1 molecules form 72 pentameric structures, or pentons, that form the basic building block (or capsomere) of the capsid. Each penton consists of a ring of five β-barrel-containing VP1 monomers. Together these form a $T=7d$ lattice.
(Figure 3-3), with an invading arm comprised of the C-termini of each VP1 undergoing exchange with neighbouring pentons to stabilise the capsid shell (Liddington et al., 1991). Structural and biochemical studies have also highlighted the importance of calcium ions and disulphide bonds between pentons in capsid assembly (Ishizu et al., 2001; J. Nilsson et al., 2005). Although there are no X-ray structures for intact human polyomaviruses, the isolated pentons from both JCV and BKV have been solved, and have shed much light on the binding of cell surface receptors (Neu et al., 2010; Neu et al., 2013). Cryo-EM structures of intact polyomaviruses have been reported but lack the resolution to identify calcium ions, disulphides or bound receptors (excluding work presented in this thesis, current structures range from 11-25 Å (T.-C. Li et al., 2003; J. Nilsson et al., 2005; Shen et al., 2011; J.P. Griffith et al., 1992).

In all existing polyomavirus X-ray structures, the N-terminal 13-15 residues of the VP1 subunits were not crystallographically resolved, which was proposed to result from the disorder in this region (Gillock et al., 1997). Subsequent biochemical studies demonstrated that these residues contain a DNA-binding domain (DBD) (P.P. Li et al., 2001; Gillock et al., 1998). When expressed alone, VP1 can self-assemble into VLPs and package dsDNA molecules of comparable size to the viral genome (Gillock et al., 1997). It has therefore been proposed that this N-terminal DBD extend towards the centre of the native particle to interact with the encapsidated ~5.2 kB dsDNA genome (Liddington et al., 1991), but this has not been directly visualised.

### 3.1.2.2 Minor capsid proteins

The minor capsid proteins VP2 and VP3 are not resolved in existing X-ray crystallography structures of SV40 and PyV. VP2/3 are expressed from the same open reading frame; however VP2 has an extended N-terminus, which contains a site for myristoylation (Krauzewicz et al., 1990). A single copy of VP2/3 has been proposed to bind within a cavity on the internal face of each VP1 penton. It has been hypothesised this occurs in one of five symmetry-related orientations, consistent with previous difference maps calculated from very low-resolution X-ray diffraction data (J.P. Griffith et al., 1992). The binding of the common C-terminal 105 residues of VP2/3 in this location has been observed in PyV penton (X.S. Chen et al., 1998), although only ~20 amino acids were crystallographically ordered and resolved at high-resolution. This region of VP2/3 has been proposed
to insert in a hairpin-like manner into the cavity of the VP1 penton, where it binds via hydrophobic interactions. However, much of the VP2/3 sequence is thought to be intrinsically disordered (X.S. Chen et al., 1998), with the common C-terminal region containing an additional DBD to that found in VP1. In SV40, VP2/3 are essential for viral morphogenesis, as shown by mutagenesis experiments (Clever et al., 1993; Dean et al., 1995). However, direct interactions between the minor capsid proteins with a packaged polyomavirus genome have not been visualised structurally. Both SV40 and BKV VP2/3 proteins also contain a C-terminal nuclear localisation signal (NLS). Mutations of basic residues in this region have been shown to decrease nuclear entry during host cell infection (Nakanishi et al., 2002; Nakanishi et al., 2007; Bennett et al., 2015). The functional indispensability of the minor capsid proteins makes them an attractive target for antiviral therapy.

Figure 3-3 – Polyomavirus structure. Polyomavirus capsids are composed of 72 pentamers of the major capsid protein VP1 in a T=7d arrangement. Twelve pentamers (white) are situated on five-fold symmetry axes, whereas the sixty others (6 of which are shown, coloured) are at 6-coordinated positions. The VP1 C-terminal arm “invades” a neighbouring pentamer, in the pattern shown, forming a beta strand in a “receiving” β-sheet. (Adapted from (Wolf et al., 2010)).
3.1.2.3 Genome packaging

Polyomavirus genomes are encapsidated with the cellular histone proteins H2A, H2B, H3 and H4 (Pagano, 1984), forming a chromatin-like structure termed a ‘minichromosome’; early work suggested it contains 20-24 nucleosomes (Cremisi et al., 1975; J.D. Griffith, 1975; U. Müller et al., 1978). However, unlike cellular chromatin, polyomavirus minichromosomes lack histone H1 (Fang et al., 2010), which is thought to be required for chromatin compaction (Thoma et al., 1979). This observation implies that the minichromosome is not highly compacted within the virus particle. Early attempts to capture the structure of isolated SV40 minichromosomes were carried out using cryo-electron microscopy (Dubochet et al., 1986). These experiments revealed that the isolated genome adopts a range of structures reminiscent of cellular chromatin when outside the capsid and that the structure observed varies depending on pH (Christiansen et al., 1977) and salt concentration (J.D. Griffith, 1975). However, no structural information for the minichromosome, in situ, has been reported, although small angle X-ray scattering and coarse-grained computer simulations suggest that packaged nucleosomes lack a highly ordered structure, but are bound at the inside surface of the capsid layer (Saper et al., 2013). Unlike other dsDNA viruses (e.g., herpes simplex virus (HSV)) (D.W. Bauer et al., 2015), polyomaviruses do not utilise powerful packaging motors to encapsidated their genome, suggesting that assembly occurs spontaneously or requires a different manner of protein-DNA interaction.

3.1.3 The BKV lifecycle

3.1.3.1 Attachment

BKV infection begins with attachment of the virion to the host cell surface. Studies have shown that BKV genotype Ia can utilise the b-series gangliosides GD3, GD2, GD1b and GT1b as cellular receptors - Figure 3-4 (Neu et al., 2013; Low et al., 2006). Crystallographic soaking experiments with isolated BKV pentons and GD3 oligosaccharide demonstrated that an α(2,8)-disialic acid motif is the minimal binding epitope present on these molecules. The presence of ‘left arm’ (Figure 3-4) on the major BKV receptors GD1b and GT1b enhances infection over the smaller counterparts, but the molecular basis of this enhancement remains uncertain as these cannot be soaked into penton crystals. Different BKV
genotypes represent distinct serotypes and may utilise a different spectrum of cell surface receptors, resulting in an altered tropism (Pastrana et al., 2013). It has also been reported that BKV may be able to recognise an α(2,3)-linked sialic acid present on an N-linked glycoprotein (Dugan et al., 2005). In addition, recent reports have shown that BKV VLPs, but not isolated pentons, can use glycosaminoglycans (GAGs) to attach to target cells (Geoghegan et al., 2017). This observation suggests that GAGs bind to regions that are not presented on pentons, presumably where different capsomeres interact. Therefore, studies to identify the GAG binding site on BKV will need to be performed on intact capsids.

Figure 3-4 – Ganglioside receptors of BKV. Schematic representation of b-series gangliosides used by BKV. (Adapted from (Neu et al., 2013).)

3.1.3.2 Internalisation

It has been shown that following attachment to Vero cells, BKV can utilise caveola-mediated endocytosis (Eash et al., 2004). This process has also been demonstrated in physiologically-relevant renal proximal tubular epithelial (RPTE) cells. Between 2-4 hours post infection, BKV virions reach neutralising antibody resistant compartments and peak colocalisation of BKV with Cav-1 is observed after 4 hours (Moriyama et al., 2007; Moriyama and Sorokin, 2008b). However, a recent study by Zhao et al. demonstrated entry of BKV into RPTE cells by a caveolin- and clathrin-independent pathway (Zhao et al., 2016). In light of this anomalous result, further studies are required to determine the precise mechanism BKV uses to enter cells.
3.1.3.3 Trafficking through the Endoplasmic Reticulum

Following internalisation, BKV is thought to enter endosomes. Studies using lysosomotropic agents such as ammonium chloride have demonstrated that acidification of endosomes ~2 hours post-entry is essential for BKV infection (Eash et al., 2004; M. Jiang et al., 2009). Intracellular trafficking of BKV requires an intact microtubule network and studies in RPTE cells have demonstrated that microtubule dynamics are essential (Eash and Atwood, 2005; Moriyama and Sorokin, 2008a; M. Jiang et al., 2009). Colocalisation of BKV with markers of the Golgi network have however not been observed, supporting the idea that the Golgi apparatus is bypassed during infection (Low et al., 2006; Moriyama and Sorokin, 2008a; M. Jiang et al., 2009; Nelson et al., 2013). Using a genome-wide, small interfering RNA (siRNA) screen, Zhao et al. recently identified factors involved in trafficking of BKV from late endosomes to the ER. Their results demonstrated that following endosomal sorting, BKV is transported in Rab18-positive vesicles which are subsequently captured by the NRZ complex which is involved in retrograde transport to the endoplasmic reticulum (ER) (Zhao and Imperiale, 2017). BKV is thought to reach the ER at approximately ~10 hours post-entry, as evidenced by colocalisation studies using brefeldin A, an inhibitor of protein transport, or observing colocalisation with markers for the ER (Moriyama and Sorokin, 2008a; M. Jiang et al., 2009). Autophagy has also been implicated in entry/intracellular trafficking of BKV. When cells are treated with inhibitors of autophagy during the first 8 hours of the BKV lifecycle, infection is reduced. A similar effect is observed in cells which under express genes involved in autophagy such as Beclin-1 and ATG7 (Bouley et al., 2014).

3.1.3.4 Release from the ER and Nuclear Entry

An interesting feature of polyomaviruses is their requirement to travel through the ER during cellular entry, before reaching the cytosol. Trafficking through the ER exposes BKV to reducing conditions, isomerases and ER resident chaperones. These host factors appear to cooperatively promote partial uncoating of the virion and render permissive for ER to cytosol transport. It has been shown that trafficking through the ER influences disulphide bond isomerisation and the protease cleavage pattern of the major capsid protein VP1 (M. Jiang et al., 2009). The ER resident reductase, ERdj5, is known to reduce disulphide bonds present in the BKV capsid and by doing so, promote BKV infection (Inoue et al., 2015). A
number of host factors involved in the ER-associated protein degradation (ERAD) pathway and the proteasome have been implicated in the transport of BKV from the ER to the cytosol (Dupzyk and Tsai, 2016). Inhibition of proteasomes or the ERAD pathway with epoxomicin and eeyarestatin, respectively, results in reduced BKV infection and accumulation of virions in the ER with increased exposed of VP2/VP3 (Bennett et al., 2013). Externalised VP2 is proposed to create a hydrophobic surface which can bind to and integrate with the ER membrane, resulting in the release of partially uncoated virions into the cytosol (Bennett et al., 2013; Nelson et al., 2013). Before membrane penetration, ER-resident Hsp70 BiP may be recruited to the hydrophobic virions to maintain them in a soluble state. The release from BiP is thought to be triggered by Grp170 which converts ADP-BiP to ATP-BiP. Once embedded in the ER membrane, BKV is believed to induce reorganisation of ER membrane proteins by promoting recruitment of the J-proteins B12 and B14. These proteins recruit the Hsc70 chaperone family via their cytosolic domain and cooperatively eject the virus particle into the cytosol (Inoue and Tsai, 2015; Goodwin et al., 2011). In addition, the major capsid protein, VP1, has been reported to interact with ER membrane resident Derlin-1 which mediates penetration of the ER membrane (M. Jiang et al., 2009). Subsequently, Hsp105 and SGTA proteins can be recruited to extract BKV into the cytosol (Ravindran et al., 2015; Walczak et al., 2014).

Once BKV reaches the cytosol, it must be transported into the nucleus so it can be replicated. Studies have demonstrated that exposed NLS signals on VP2 and VP3 can be recognised by the importin α/β1 import pathway which allows the viral genome to enter the nucleus through the nuclear pore complex (Bennett et al., 2015).

3.1.3.5 Assembly and release

Concurrently with genome replication (see 1.1.1) translation of the structural proteins VP1, VP2 and VP3 occurs in the cytoplasm, before subsequent transport to the nucleus for virion assembly. This transportation process is mediated by an NLS signal on the N-terminus of VP1 (Bennett et al., 2015). Studies have shown that chaperone proteins can associate with the C-terminus of VP1 in the cytoplasm to prevent premature capsid formation. Once inside the nucleus, high-calcium levels promote assembly of viral capsomeres (consisting of five VP1 molecules and a single copy of VP2 or VP3), around newly synthesised viral
genomes (Chromy et al., 2003; Teunissen et al., 2013). BKV virions begin to appear in infected cells 48 hours post-infection (Low et al., 2004). Each infected cell can contain up to 6000 BKV virions, as estimated from renal biopsies from PVAN patients (Hirsch and Steiger, 2003). Dense accumulations of virions within the nucleus of infected cells have also been visualised by electron microscopy (Drachenberg et al., 2003).

The release of progeny BKV has been shown to occur through a lytic mechanism in RPTE cells, although strong cytopathic effects are less-commonly observed compared to SV40 (Low et al., 2004; Henriksen et al., 2016). Since BKV establishes lifelong persistent infections in humans, it is questionable as to whether the virus relies solely on lytic release. Recent studies in RPTE cells show that approximately 1% of progeny virions can be released by a non-lytic mechanism. This process is sensitive to inhibition with an anion channel blocker 4,4′-diisothiocyanato-2,2′-stilbenedisulfonic acid, which affects cellular secretion pathways. Also, under these conditions, BKV virions became trapped in acidic compartments derived from late endosomes and lysosomes (Evans et al., 2015). Data from Panou et al. recently demonstrated a role for the BKV agnoprotein in cellular egress. They observed that the BKV agnoprotein is required for nuclear egress of virions and that this process is likely mediated through an interaction with α-soluble N-ethylmaleimide sensitive fusion (NSF) attachment protein (α-SNAP) which is involved in vesicular trafficking (Panou et al., 2018).
Figure 3-5 – BKV replication cycle. Model of the BKV life cycle showing (1) attachment, (2) internalisation, (3) trafficking, (4) uncoating, (5) nuclear entry, (6) early gene expression, (7) DNA replication, (8) late gene expression, (9) virion assembly, (10) lytic release and (11) non-lytic release. (Figure taken from (Helle et al., 2017) [published under a creative common [CC BY 4.0] license].)

3.1.4 Epidemiology and tissue tropism

Over 70% of children are seropositive for BKV by the age of 10, but there is still debate as to the primary route of transmission (Knowles, 2006). Detection of BKV in the tonsils of children supports the idea that the respiratory route is the likely mode of transmission (Goudsmit et al., 1982), however other routes have been proposed such as faecal-oral, transplacental or via blood transfusion (Bofill-Mas et al., 2000; Taguchi et al., 1975; Chatterjee et al., 2000). Primary infection with BKV is generally asymptomatic but may occasionally be associated with a mild respiratory disease (Goudsmit et al., 1982). It is thought that BKV enters the bloodstream via infected tonsillar tissue before infecting blood mononuclear cells which allow the virus to disseminate to secondary sites within the body such as the kidney (Chatterjee et al., 2000). BKV DNA has also been detected in lymph nodes, leukocytes and the brain (Chatterjee et al., 2000; Vago et al., 1996). Once the primary infection is resolved, BKV maintains a persistent infection of kidney epithelial cells and may occasionally reactivate resulting in asymptomatic viruria.
(Doerries, 2006). The virus is occasionally shed in the urine; however, this is less common in healthy individuals than immunosuppressed patients (Bressollette-Bodin et al., 2005). Further research is required to elucidate the mechanism of viral persistence and the precise factors which lead to its reactivation in immunosuppressed individuals. Studies have shown that BKV encoded microRNA plays a vital role in limiting the replication of the archetype virus in a natural host model of infection and may thus aid in viral persistence (Broekema and Imperiale, 2013).

### 3.1.5 BKV-associated diseases

#### 3.1.5.1 Polyomavirus-associated nephropathy

Approximately 10% of kidney transplant recipients (KTRs) experience PVAN and up to 90% of these individuals will go on to lose their graft (Ramos et al., 2009). PVAN usually presents 10-13 months after renal transplantation and is characterised by inflammatory interstitial nephropathy and intranuclear viral inclusions in cells of the glomeruli, collecting ducts and renal tubules (Boothpur and Brennan, 2010; Nickeleit et al., 1999). In the final stages of PVAN, inflammation, interstitial fibrosis and tubular atrophy manifest and collectively contribute to graft loss (Hirsch et al., 2005). The first signs of PVAN are increased levels of serum creatinine, due to poor clearance of this molecule by the impaired kidney. This is followed by viruria (virus in urine) which is detectable in 35-57% of KTRs, and in 13-22% of PVAN patients, viruria progresses to viremia (virus in blood) (Hirsch et al., 2002). The gold standard method of PVAN diagnosis is a biopsy of the affected kidney. The incidence of PVAN in KTRs is increasing as a result of more potent immunosuppressive regimes used to prevent acute rejection. These usually consist of a corticosteroid, an antimetabolite and a calcineurin inhibitor (Ramos et al., 2009; Suwelack et al., 2012; Binet et al., 1999). There are several pre-transplant risk factors for the development of PVAN in KTRs which can be divided into host and graft characteristics. Risk factors for the recipient include a history of graft loss by PVAN, older age, male gender and being seronegative for BKV. Graft risk factors include ischemic or immunological injury, the degree of human leukocyte antigen mismatch and a seropositive donor. Viral factors can also influence the risk of developing PVAN. It is known that mutations in the VP1 capsid protein of BKV are associated with increased pathogenesis (Randhawa et al., 2002). A recent study reported that BKV might
utilise host APOBEC3 cytosine deaminases to obtain mutations in VP1 which alter glycan receptor usage and to evade antibody-mediated neutralisation (Peretti et al., 2018).

3.1.5.2 Haemorrhagic cystitis

HC is a painful bladder condition with symptoms including hematuria, dysuria, urgency and subarapubic pain. In 10-25% of bone marrow transplant patients, BKV can cause HC with symptoms usually presenting two weeks post-transplantation (Dropulic and Jones, 2008). Approximately 50% of bone marrow transplant patients have detectable BKV viruria, and individuals with plasma viral loads over $10^4$ copies/ml are at high risk of developing HC (Erard et al., 2005). Diagnosis of BKV-associated HC can be performed by PCR amplification of viral DNA from serum or urine, or by identification of decoy cells in the urine of HC patients. These cells exhibit enlarged nuclei and a large basophilic intranuclear inclusion (Wong, 2006). At present, BKV-associated HC treatment is limited to bladder irrigation, forced diuresis, hyperhydration and pain management (Dropulic and Jones, 2008).

3.1.5.3 Cancer

Polyoma is derived from the Greek words “poly” (many) and “oma” (tumours). Various cancers have been linked to BKV, including urothelial tumours and prostate cancer (Weinreb et al., 2006; Newton et al., 2006; Toptan et al., 2016). However, there is still much debate as to whether or not BKV is the causative agent for these malignancies. It is known that the early-coding region of BKV can induce transformation in both human and rodent cells (Trabanelli et al., 1998) and inoculation of rodents with BKV can induce tumour formation (Chenciner et al., 1980; Corallini et al., 1978). The early-encoded TAg and tAg, are responsible for driving neoplastic transformation by inactivating proteins involved in tumour suppression, such as pRb and p53 (Bocchetta et al., 2008; Topalis et al., 2013; Fioriti et al., 2003). The presence of TAg-p53 complexes has been observed in tissue obtained from prostate cancer samples (D. Das et al., 2008; D. Das et al., 2004). Furthermore, it has been demonstrated that tAg can increase activation of the mitogen-activated protein kinase pathway which may result in increased cell proliferation and transformation (Sontag et al., 1993). It has been shown that BKV DNA is more prevalent in early-stage cancer tissue when compared to healthy
controls. Such observations indicate a possible ‘hit and run’ mechanism of BKV-induced oncogenesis (D. Das et al., 2008; D. Das et al., 2004). The role of BKV in the development of human cancers is an ongoing subject of discussion. At present, BKV is classified as “possibly carcinogenic to humans” (group 2B) by the International Agency for Research on Cancer (Bouvard et al., 2012).

3.1.6 Current therapeutic strategies

In general, PVAN is managed by frequent PCR monitoring of BKV viremia, coupled with a gradual reduction in immunosuppression (Weiss et al., 2008; Alméras et al., 2011; Alméras et al., 2008; Hardinger et al., 2010). However, this strategy runs the risk of acute rejection of the transplanted kidney as a result of insufficient immunosuppression (Menter et al., 2013). Another consequence of modifying the immunosuppressive regimens of PVAN patients is an increased incidence of chronic rejection (Sawinski and Goral, 2015).

There are currently no targeted antivirals available to treat BKV-associated diseases. Several generic antivirals have been reported as having anti-BKV properties, but the majority of such reports come from uncontrolled retrospective observational studies (Ambalathingal et al., 2017). Also, most tested therapeutic agents have been used in conjunction with a gradual reduction in immunosuppression, meaning it is difficult to draw definitive conclusions about the clinical efficacy of such treatments. A summary of current strategies to treat BKV-associated diseases is included below.

3.1.6.1 Cidofovir

Cidofovir is a cytidine nucleotide analogue which inhibits viral, and to a lesser extent human, DNA polymerases (Safrin et al., 1997). There are a number of single-centre case reports which describe the clinical efficacy of the antiviral drug cidofovir when prescribed alongside a gradual reduction in immunosuppression. However, there has been no reported randomised trials thus far (Kuypers et al., 2005; Kwon et al., 2013; Muller et al., 2005; Savona et al., 2007). The major problem with cidofovir is that it is itself associated with nephrotoxicity, which often limits its clinical application. An orally administrable prodrug of cidofovir, Brincidofovir (CMX001), is currently undergoing phase III clinical trials. Brincidofovir has a lower incidence of nephrotoxicity than cidofovir and as such
holds promise as a potential anti-BKV treatment (Marty et al., 2013; Reisman et al., 2014; Papanicolaou et al., 2015).

### 3.1.6.2 Leflunomide

The immunosuppressant agent, Leflunomide, has been shown to have antiviral properties against BKV in vitro (Liacini et al., 2010). In several cases, Leflunomide has been used as part of the immunosuppressive regimen for patients affected by BKV-associated diseases (Williams et al., 2005; Teschner et al., 2009; J.K. Wu and M.T. Harris, 2008; Faguer et al., 2007; Josephson et al., 2006; Krisl et al., 2012). These studies demonstrated an association between leflunomide and a decrease in BKV viral load, but it is difficult to conclude if this is due to a direct antiviral effect or overall immunosuppression. Furthermore, there are adverse side effects of leflunomide which include bone marrow suppression, haemolysis, thrombotic microangiopathy and hepatitis (Muller et al., 2005). Controlled studies must be carried out to confirm the safety and efficacy of this drug as a treatment for BKV-associated diseases (Krisl et al., 2012; Cuellar-Rodriguez et al., 2013).

### 3.1.6.3 Fluoroquinolones

Fluoroquinolone antibiotics have been shown to have an inhibitory effect on BKV and SV40 in vitro and therefore represent a potential therapy for patients suffering from BKV-associated diseases. In vitro studies have shown that fluoroquinolones can reduce expression of TAg and also inhibit its helicase activity (Sharma et al., 2011; Ali et al., 2007; Portolani et al., 1988). Evidence of clinical efficacy comes from retrospective analysis of a clinical trial where fluoroquinolones were administered as antibiotic prophylaxis at the time of kidney transplantation. The results demonstrated that people treated with a fluoroquinolone had a reduced chance of developing BKV viremia (Wojciechowski et al., 2012; Gabardi et al., 2010). Unfortunately, conflicting results were published from two randomised studies of KTRs receiving the fluoroquinolone, levofloxacin, which showed no clinical efficacy (B.T. Lee et al., 2014; Knoll et al., 2014). Together, these results do not demonstrate a convincing therapeutic benefit of fluoroquinolone treatment for BKV-associated diseases (Hirsch et al., 2013).
3.1.6.4 Neutralising antibodies

In the last few years, there has been increasing interest in the possibility that BKV nephropathy might be prevented or treated using antibodies that neutralise various strains of the virus (Pastrana et al., 2012; Randhawa et al., 2015). Such antibodies could either be administered as recombinant mAbs or induced through vaccination. The first human mAbs capable of binding BKV virions were reported in late 2015 (Jelcic et al., 2015). It is not yet clear which epitopes these potentially useful new mAbs bind and as such, there is a pressing clinical need to better understand the molecular biology of antibody-mediated neutralisation of BKV (Buck, 2016). Such effort will be bolstered by the availability of high-resolution information for BKV-receptor/mAb complexes and robust pipelines to determine such structures. Patents for polyomavirus neutralising antibodies and immunogenic compositions were filed in 2016 and 2017 respectively (Abend et al., 2018; Buck et al., 2017). Furthermore, a phase II clinical trial to determine the safety, tolerability and efficacy of a BKV-specific, pan-serotype neutralising antibody (MAU868) is set to commence on October 15th, 2018 (ClinicalTrials.gov Identifier: NCT03456999).

3.1.7 Project aims

Since the early crystallographic studies of SV40 and PyV, there has been a paucity of high-resolution structures of intact human polyomaviruses. Whilst low-resolution cryo-EM maps of intact particles have been reported, the resolution achievable has been limited to worse than one nanometre. In contrast, very high-resolution structures of isolated pentons can be obtained, but are limited as to what they can reveal about the native virion. Given the growing interest in using neutralising mAbs to treat BKV-associated diseases, the ability to produce detailed structures of this virus would be useful for determining their mechanism of binding. This project aims to establish a robust cryo-EM workflow for determining high-resolution structures of BKV, and by doing so, attempt to address several unanswered questions regarding this clinically important human pathogen, namely:

- How is the viral genome packaged?
- Does the BKV capsid differ to SV40?
- How does BKV recognise GT1b ganglioside?
• Where is the glycosaminoglycan binding site?

In more general terms, this project aims to demonstrate the utility of cryo-EM for studying glycan binding to viruses at resolution better than 4 Å.

This chapter describes the results achieved towards answering the above aims. All work presented was performed by the candidate unless stated otherwise.
3.2 Results part one: Intermediate-resolution studies of BKV

In this section, the first sub-nanometre resolution solution structures of the native, infectious BKV virion, and a BKV VP1-only VLP are reported. Together, these structures unambiguously identify the location of VP2/3 within the capsid of a human polyomavirus. We also identify discrete bridges of electron density, which connect the VP1 N-terminus to packaged dsDNA and observe points of contact between density for the minor capsid proteins and the viral genome. Furthermore, we see shells of density in the centre of the virion that provide structural evidence about how the minichromosome could be packaged.

3.2.1 Production of infectious BKV

To produce virions for structural studies, the circularised Dunlop strain genome of BKV was transfected into Vero cells, a cell line derived from monkey kidney epithelial cells, which is ideal for culturing the virus. Ten days post transfection, cells were lysed, and the resulting crude cell lysate containing BKV virions was used to infect naïve Vero cells, which were grown for a further 14 days. Western blotting analysis showed that the major structural protein VP1 was present in both the media and crude cell lysate (Figure 3-6A), demonstrating that BKV replicates in Vero cells, and suggesting that much of the virus made is not released into the media but remains inside cells. Virions were then purified as described previously (M. Jiang et al., 2009), by centrifugation through a caesium chloride gradient.

To confirm that this material was infectious, an aliquot of the purified virus was used to infect RPTE cells, which were analysed three days post infection by Western blotting using antibodies against VP1 and VP2/3. All three structural proteins were present in the virus particles (Figure 3-6B). The infected cells were also analysed by quantitative PCR using primers against the BKV genome, which gave a genome copy number of ~1x10^8 per μg of total cell-extracted DNA, verifying that the infection was productive as both genome replication and late gene expression had occurred (data not shown). RPTE cells grown on glass coverslips were also analysed by immunofluorescence microscopy three days post infection with antibodies against VP1 and VP2/VP3 (Figure 3-6C). Whilst low-level diffuse staining for the capsid proteins can be observed throughout the cytoplasm, the structural proteins are primarily localised in the nucleus, with VP1 staining appearing as distinct puncta. Negative stain electron microscopy (EM)
images (Figure 3-6D) confirmed the presence of polyhedral particles with a diameter of 45-50 nm in the crude cell lysate, together with copious cellular debris. However similar images of the CsCl-gradient purified virus showed that although considerably purer than the crude cell lysate, significant low molecular weight contaminants remained (Figure 3-6E). An additional purification step of centrifuging through a 1 MDa spin concentrator was therefore used. The virus, which was retained by the filter whilst contaminants washed through, was essentially pure after this step (Figure 3-6F) and taken forward for cryo-electron microscopy (Figure 3-6G).
Figure 3-6 – Purified BKV virions are infectious. (A) Vero cell lysate and 30 μl of growth media were analysed by western blotting to determine the presence of VP1 capsid protein and GAPDH expression. (B) The lysate from RPTE cells infected by gradient purified virions was separated by SDS-PAGE and analysed by western blotting for the VP1 and VP2/VP3 capsid proteins. (C) RPTE cells were grown on glass coverslips and infected with virions; 3-dpi slides were fixed with 4% paraformaldehyde and stained with antibodies against VP1 (green) and VP2/VP3 (red). The nuclei were visualised with DAPI (blue). Scale bars, 10 μm. (D) Negative-stain electron micrograph of virions present in the Vero cell lysate. (E) Partially purified virions following centrifugation through a caesium chloride gradient. (F) Purified virions following final purification step through a centrifugal concentrator (right). (G) Purified virions suspended in a layer of vitreous ice following cryo-grid preparation. Scale bars, 100 nm. (Figure taken from Hurdiss et al., 2016) [published under creative commons {CC BY} license].) Immunofluorescence experiments (C) were performed by Dr Emma Prescott and Margarita Panou, University of Leeds.
3.2.2 Production of VP1 only virus-like particles

To produce VLPs for comparison to the native virion, HEK293TT cells were transfected using a codon-optimised VP1 expression plasmid, together with a reporter plasmid encoding eGFP. The resulting VLPs were purified on an OptiPrep® gradient and the presence of VP1 detected via Western blotting (Figure 3-7A). The purified VLPs were used to transduce naïve HEK293TT cells, with the resulting fluorescence positive cells demonstrating the VLP’s ability to both package the eGFP reporter plasmid and deliver it across cell membranes (Figure 3-7B). For cryo-EM studies, VLPs were concentrated in the same manner as the virions. Negative stain (Figure 3-7C) and cryo-EM images (Figure 3-7D) of VLPs demonstrated that they have a size and morphology indistinguishable from the native virion.

**Figure 3-7 – BK polyomavirus VLP production in HEK293TT cells.** (A) Western blot for VP1 expression in each fraction from the iodixonal gradient. 20 µl of the gradient solution was loaded and VP1 levels were detected using a VP1-specific antibody. (B) Purified VLPs were used to transduce HEK293TT, and the presence of EGFP was analysed by light microscopy. The image is representative and approximately 65% of the cells were transduced. (C) Negative-stain micrograph of purified VLPs. (D) Cryo-electron micrograph of purified VLPs suspended in vitreous ice. Scale bars, 100 nm. (Figure taken from Hurdiss et al., 2016) [published under creative commons (CC BY) license]. VLP expression, purification and transduction experiments (A-B) performed by Ethan Morgan, University of Leeds.
The VLP and virions were further characterised by assessment of their histone and DNA content (Figure 3-8). Each type of capsid packages genomic DNA, but the virion contains the BKV genome, and more histones, whilst the VLP appears to package the high copy number eGFP reporter plasmid and a lower level of histone proteins. These observations are in agreement with previous studies which have demonstrated the ability of SV40 and PyV VP1 protein to promiscuously package cellular DNA and histones (Gillock et al., 1997; Trilling and Axelrod, 1970). The presence of encapsidated histones was also investigated by mass spectrometry analysis of purified virions which identified low molecular weight species consistent with histone proteins H2A, H2B, H3 and H4 (data not shown).

Figure 3-8 – VLPs and virions contain cellular histones and DNA. (a) Purified VLPs and virions were boiled in gel loading buffer and analysed by western blot for histone content using an anti-histone H3 antibody VP1 expression was used as a loading control and VP2/3 expression was used to confirm minor capsid protein expression in the virions alone. (b) BamHI digested and undigested DNA extracted from VLPs and virions was analysed on a 0.7% agarose gel. BamHI linearises the pEGFP reporter plasmid/viral genome and digests cellular DNA, shown by the presence of a single discrete band in the digested samples. Linearised pEGFP and pGEM7 BKV genome were included as controls. [Figure taken from (Hurdis et al., 2016)] Experiments performed by Ethan Morgan, University of Leeds.
3.2.3 The sub-nanometre structure of BKV and VLP

Images of unstained, frozen-hydrated virions and VLPs were recorded using an electron microscope fitted with a direct electron detecting camera, allowing the solution structures of the virion and VLP to be resolved at 7.6 Å and 9.1 Å respectively (Figure 3-9A). These are the first sub-nanometre resolution EM structures for any polyomavirus capsid (T.-C. Li et al., 2003; J. Nilsson et al., 2005; Shen et al., 2011) and allowed visualisation of secondary structural elements (α-helices and β-sheets, although not individual β-strands). Both the native virion and VLP are isometric particles with a diameter of ~500 Å, similar to that described previously for other polyomaviruses, including SV40 (Liddington et al., 1991; Stehle et al., 1996) and MPV (Stehle and Harrison, 1996). The capsids have a $T=7d$ quasi symmetry; and are built from 72 pentameric capsomers which are easily visible in the maps. Each penton contains five copies of the major structural protein VP1, and there are six distinct conformations of VP1 in the shell (Figure 3-9B). The pentamers are tied together using C-terminal arms, with each pentamer of VP1 donating and receiving 5 such arms to/from adjacent pentamers (Figure 3-9C). A homology model of the BKV VP1 asymmetric unit was generated using the SWISS model server and flexibly fitted into the corresponding density of the virion map (Figure 3-10). At the resolution presented the gross fold of BKV VP1 is extremely similar to previous structures of the non-human polyomavirus, SV40, which shares 81.7% sequence identity in the major structural protein.
Figure 3-9 – The cryo-EM structures of BKV and VLP. (A) An external view of the virion (left) and VLP (right) shown at contour levels of 0.022 and 0.009, respectively. The electron density maps have been sharpened using a negative B factor correction of \( B = -456 \) and \(-804 \) Å²), and the density is coloured using a radial scheme shown in Ångströms. (B) The architecture of a polyomavirus capsid showing how the \( T = 7d \) capsid is built from 72 pentamers of VP1, and an identical VP1 polypeptide sequence is found in six distinct quasi-equivalent conformations in the capsid shell (1, red; 2, yellow; 3, green; 4, cyan; 5, blue; and 6, grey). (C) The cryo-EM structure of the virion (left) and VLP (right) coloured according to the scheme in (B), showing the arrangement of different VP1 molecules in three dimensions. (Figure taken from (Hurdiss et al., 2016) [published under creative commons {CC BY} license].)
Figure 3-10 – The BKV homology model. A 20-Å thick slab through the BKV homology model which is fitted as a rigid body in the sharpened virion map. The EM density is shown as a blue mesh at a contour level of 0.0193 and the six quasi-equivalent chains are coloured as in Figure 3-9.

3.2.4 Minor Capsid Proteins and the organisation of the dsDNA genome

Unlike previously described X-ray structures, the virion map contains significant density that were attributed to the minor capsid proteins VP2 and VP3 (Figure 3-11A). Low-resolution data from related viruses has suggested that VP2 and VP3 bind to the inner surface of a VP1 penton, but this has not been previously visualised in BKV. This density is weaker (and at a lower resolution) than the VP1 shell but can still be clearly seen even in a map which has been subjected to the B-factor correction used to reveal high resolution features, the unsharpened/unmasked maps for the virion and VLP (9.1 Å and 11.07 Å respectively) are deposited under same EMDB codes as their refined counterparts. A depression on the inside surface of the VP1 penton appears to
be completely filled with a conical density, which was attributed to VP2 and/or VP3 on the basis that this is missing in the VLP structure (Figure 3-11B).

Both the virion and the VLP have additional density inside the capsid shell, which were proposed to arise from the packaged dsDNA genome, and eGFP reporter plasmid/cellular DNA respectively. The virion map has strong density in two distinct shells directly beneath the capsid layer. The outermost of these is connected to the proposed VP2/3 density and the two shells have a thickness and radial spacing consistent with dsDNA wrapped around a histone octamer (Figure 3-11B) (Luger et al., 1997). Such discrete bands are reminiscent of those observed in projection images of nucleosomes (Bilokapic et al., 2018). The VLP map lacks these two discrete shells, and overall the density is weaker and more diffuse (Figure 3-11D). In the virion map, discrete bridges of density were observed between the VP1 capsid and encapsidated dsDNA, the location of these relative to the virus capsid and minor capsid proteins is shown in Figure 3-12A. These bridges are situated beneath the N-termini of each of the six VP1 quasi-equivalent conformers using the nomenclature adopted by Stehle et al – Figure 3-12B-G (Stehle et al., 1996). Interestingly, the density associated with the N-termini of chains 3 and 6 appear to be a weaker feature within the virion map. No equivalent features are visible in the VLP map. It was postulated that these densities correspond to some portion of the 13-15 N-terminal residues of VP1, which are not resolved in previous polyomavirus structures.
Figure 3-11 – Minor capsid proteins and genome organisation. (A) A 40-Å thick slab through the unsharpened/unmasked virion map shown at a contour level of 0.0034. Pyramidal density below each VP1 penton and two shells of electron density adjacent to the inner capsid layer can be seen. The density within 6 Å of the fitted coordinates for SV40 VP1 is coloured grey. The remaining density is coloured in a radial colour scheme. Density for VP2 and VP3 is coloured blue→green, and for packaged dsDNA yellow→pink. (B) Enlarged view of the pyramidal density beneath a single VP1 penton of the virion shown at a contour level of 0.0032. Strands of dsDNA wrapped around a human histone octamer (PDB: 1AOI) are shown, indicating that the two shells of density have a comparable spacing. Discrete connective density between the pyramidal density and internal shells is also apparent. Scale bars shown. (C) 2D class averages of nucleosome core particle pairs (Bilokapic et al., 2018). (D) A 40-Å thick slab through the unsharpened/unmasked VLP map shown at a contour level of 0.0011 which shows that density for the minor capsid proteins and two shells of electron density is absent. (Figure adapted from (Hurdiss et al., 2016) and (Bilokapic et al., 2018) [published under creative commons {CC BY} and {CC BY 4.0} licenses, respectively.)
Figure 3-12 – The N-terminus of BKV VP1. (A) The location of VP1 N-termini binding to the viral genome relative to the minor capsid proteins, viewed down the capsid three-fold axis of symmetry (chain colour key shown). For visual clarity, density within 6 Å of the fitted coordinates for SV40 VP1 has been removed. A single asymmetric unit from the SV40 atomic model, coloured according to the schematic shown in Figure 3-9B, was fitted as a rigid body into the unsharpened/unmasked virion map. The map was then segmented within 15 Å of the fitted model. (B–G) 20-Å thick slabs of density through the segmented map showing connective bridges of density between the capsid and packaged genome. Each of these is located beneath the N-termini of each of the six VP1 quasi-equivalent conformers, which are denoted by asterisks (map contour levels shown). (Figures taken from (Hurdiss et al., 2016) [published under creative commons {CC BY} license].)
3.3 Discussion part one

The structure of the wild-type, infectious BKV virion presented here provides new insights into polyomavirus biology. The structure and arrangement of BKV VP1 is remarkably similar to that of SV40, with minor rearrangements that do not affect either the gross fold of VP1, size of the penton body or resulting icosahedral particle. The icosahedral averaging applied to the virion structure during refinement means that information for the VP2 and VP3 proteins is compromised. These are asymmetric proteins bound into a pentameric capsomer, presumably at one of five redundant binding sites. This capsomer then sits at both hexavalent and pentavalent positions within the $T=7d$ lattice. Thus, whilst the VP1 component of the structure is appropriately averaged, VP2 and VP3 (and their modes of binding) are incorrectly averaged, and thus the density observed is hard to interpret. However, some new conclusions can be drawn. The volume of the cone-shaped density corresponding to VP2/3 is $15-18,000 \text{ Å}^3$ (at the contour shown in Figure 3-11B). Although the errors in such calculations can be considerable, these volumes are consistent with $\sim32-38\%$ of a VP2 and 45-55% of a VP3 molecule being present within the density shown (VP2=38.3kDa; VP3=26.7kDa, density of protein = 1.37 g.cm$^{-3}$). Discrete bridges from this VP2/3 density to an inner shell were observed and were ascribed to a mixture of the remainder of VP2/3, the encapsidated dsDNA genome and packaged histone proteins. There is little evidence to suggest that the packaged genome is icosahedrally-ordered, meaning that the genome information is also obscured as a result of the symmetry imposed during image processing. There is also absolutely no suggestion from classification of the image data that the packaged chromosome is condensed into a solenoid-type structure within the virion, regardless of whether such structures may be possible for isolated minichromosomes (Dubochet et al., 1986). This is consistent with the absence of histone H1 from the virus. Despite the uncertainty caused by icosahedral averaging, the radial distribution of the density is informative, as seen in the encapsidated genomes from other dsDNA viruses (Lander et al., 2013; W. Jiang et al., 2006; Cerritelli et al., 1997). Strong density is observed inside the virion, as two $\sim24\text{ Å}$ thick radial shells separated by a gap of $\sim28\text{ Å}$. The spacing of the radial density shells in the virion structure closely matches the spacing of dsDNA present within a human nucleosome (PDB 1EQZ) (Luger et al., 1997). There is
no clear density for the histones themselves, although at lower contour levels, the two layers begin to merge, presumably because of the contribution of density from disordered histones proteins. This is consistent with observations on other histone-containing complexes, such as a retroviral intasome complex (Maskell et al., 2015), where the ordered polyphosphate backbone of dsDNA is a particularly strong feature in density maps. These observations are consistent with simulations performed on SV40, which indicated that nucleosomes within the capsid centre lack orientational order, whilst those in the layer adjacent to the capsid wall, align with the boundary (Saper et al., 2013). Any structural information from the relatively disordered nucleosomes within the capsid centre appears to be averaged out completely.

Although these are the first observations of genome packaging in a polyomavirus, density for genomic material has been observed in a number of other viruses by cryo-EM e.g. Ljungan virus (Zhu et al., 2015) For ssRNA viruses including H. circularisquama RNA virus (HcRNAV; (Miller et al., 2011)) and Turnip Crinkle virus (TCV; (Bakker et al., 2012)), a similar double-shelled pattern has been observed. This probably reflects the base-pairing of single-stranded RNA to form a collapsed, partly double-stranded substrate for packaging, and the effect of basic RNA-binding domains (or “arms”) from the viral coat proteins in condensing the DNA. For a double-stranded DNA virus, whose genome is much stiffer than a partly base-paired RNA, the outcome of packaging appears to be similar, with multiple shells of dsDNA packaged into many viruses albeit with the help of powerful packaging motors (W. Jiang et al., 2006; Lander et al., 2013). The density observed in the BKV virion again has a similar size and spacing despite the lacks of any packaging motor. Perhaps the role of histones in genome packaging for small dsDNA viruses is to exploit the host’s own strategy for DNA compaction, overcoming the need for a packaging motor.

Unlike previous polyomavirus structures, discrete bridges of density were observed connecting the encapsidated dsDNA and VP1 in the virion map. This density is located beneath the N-termini of the fitted SV40 crystal structure, indicating that it corresponds to the 13-15 residues not resolved in previous X-ray structures. Although the DNA binding properties of the VP1 N-terminus have been described biochemically, this is the first visualisation of direct interaction between a polyomavirus capsid and its packaged genome. It is
unclear whether ongoing structural work may be able to resolve these interactions at high-resolution. These bridges of density suggest a role in genome packaging/recognition and/or capsid assembly, but the N-termini also facilitate genome targeting to the nucleus. Presumably this would be during viral uncoating, when the nuclear localisation signals contained in these sequences become exposed.

The structures presented here are the first sub-nanometre resolution structures of a native, human polyomavirus, and give insight into the organisation of the minor capsid proteins, their interaction with packaged genome, and the organisation of that genome within the native particle. It therefore provides the platform for future efforts to image the particle at higher resolution and without symmetry averaging as it goes through its lifecycle.

3.4 Results part two: High-resolution studies of BKV and receptor complexes

In the following section, previous intermediate resolution observations from (Hurdiss et al., 2016) are extended and the high-resolution cryo-EM structure of BKV at 3.8 Å resolution is reported. This resolution allows differences between this human pathogen and those that infect simian and murine hosts to be visualised, which includes differences in how both the C-terminal arms and disulphide bonding stabilise the capsid. The importance of disulphide bonds is confirmed by a lower-resolution structure of BKV in reducing conditions where density ascribed to disulphide bonds disappears. Furthermore, using a simple and robust method analogous to crystallographic soaking, the structures of BKV bound to the ganglioside GT1b (at 3.4 Å) and heparin (at 3.6 Å resolution) are reported. Neither receptor causes a conformational change in the capsid. Rather they function cooperatively to increase the avidity of attachment to the host cell surface. In addition to providing basic structural information about this medically important virus, this study lays the groundwork for understanding antibody-based therapies and, conceivably, small molecule inhibitors of infection or assembly.

3.4.1 Large-scale production of BKV

To improve the quality and quantity of the BKV preparation, an alternate purification strategy was implemented using a modified version of a previously
described method (Shen et al., 2011). Growth of BKV in cell culture was scaled up to 36 x T175 flasks of Vero cells and the time from infection to harvesting was increased from 14 to 21 days. During purification, an isopycnic caesium chloride gradient was used in place of the rate zonal alternative used previously (Hurdiss et al., 2016). A clear band was visible after ultracentrifugation (Figure 3-13A) and examination of the extracted material by negative stain EM revealed pure and monodisperse BKV virions which required no additional ‘polishing’ steps to remove cellular debris (Figure 3-13B). Unlike BKV purified previously (Hurdiss et al., 2016), less stain penetration of particles was observed, indicating that this purification strategy better preserves capsid integrity. Immunofluorescence visualisation of VP1 in RPTE cells was used to demonstrate the infectivity of BKV virions obtained from this alternative purification strategy (Figure 3-13C). Furthermore, this alternative method was successfully used to purify mutant BKV for a separate study investigating the role of the virally encoded agnoprotein in virion egress (Panou et al., 2018).

![Figure 3-13 – Large-scale production of BKV. (A) Isopycnic caesium chloride gradient containing a BKV positive band which is indicated by a red arrow. (B) Negative-stain micrograph of purified BKV virions – scale bar 100 nm. (C) VP1 positive nuclei present in RPTE cells infected with purified BKV.](image)

3.4.2 High-resolution structures of BKV: receptor complexes

BKV is difficult to concentrate to the levels required for optimal cryo-EM data collection, so to increase the particle concentration and distribution of BKV for cryo-EM studies, multiple aliquots of purified virions were applied to lacey carbon grids overlaid with a layer of continuous carbon (Figure 3-14A). Owing to the low concentration of the virus sample (0.5 mg/ml) and low affinity of polyomavirus-
receptor interactions (~ 1-5 mM) (Neu et al., 2008; Stehle and Harrison, 1996), to study the BKV-GT1b structure, an on-grid soaking protocol was implemented (Drulyte et al., 2018). A 20 mM solution of the oligosaccharide portion of GT1b was applied to the carbon-immobilised virions for 30s prior to blotting and plunge freezing (Figure 3-14B). Collectively, these approaches allowed the structure of native BKV and BKV-GT1b to be determined at 3.8 Å resolution (Figure 3-16A, Figure 3-14A), and 3.4 Å resolution respectively (Figure 3-16B, Figure 3-14B). The BKV and BKV-GT1b structures are essentially identical, other than the presence of extra, high-resolution density on the capsid surface corresponding to the oligosaccharide portion of GT1b (see below). The improved quality of the purified BKV sample likely facilitated the higher resolution obtained in the resulting cryo-EM structures compared to previous studies (Hurdiss et al., 2016).

As expected, the BKV capsid has a $T=7d$ quasi-symmetry and is built from 72 pentons of the major capsid protein VP1 (Figure 3-19A). Each of the six distinct conformations of VP1 present within the asymmetric unit of the capsid shares a common core fold (Figure 3-19B/C), which is the $\beta$-sandwich, 'jelly-roll' fold, characteristic of many viral capsid proteins. The C-terminal arms which invade adjacent capsomers to build the capsid, are well-resolved, and at a slightly lower resolution than the VP1 core (on average ~0.2 Å poorer; Figure 3-17 and Figure 3-18).

**Figure 3-14 – Micrographs from BKV and BKV-GT1b data collection.** Typical micrographs from unliganded BKV (A) and BKV-GT1b (B) data sets (scale bar = 50 nm) (Figure adapted from (Hurdiss et al., 2018) [published under creative commons {CC BY} license]).
Figure 3-15: Fourier shell correlations (FSC) for BKV and BKV-GT1b. Based on the 0.143 criterion for the ‘gold standard’ comparison of two independent data sets, the resolution of the reconstructions is 3.8 Å (A) and 3.4 Å (B), respectively. (Figure adapted from (Hurdiss et al., 2018) [published under creative commons (CC BY) license]).
Figure 3.16 – Structures of BKV and BKV-GT1b. (A) Isosurface representation of the 3.8 Å structure of BKV viewed down the icosahedral two-fold and coloured according to the radial colouring scheme shown. (B) Isosurface representation of the 3.4 Å structure of BKV, in complex with GT1b (magenta). Representative EM density containing the refined atomic model is shown below each respective structure to highlight the quality of the maps. (Figure taken from (Hurdiss et al., 2018) [published under creative commons {CC BY} license])
Figure 3-17 – Local resolution of BKV and BKV-GT1b maps. Local resolution of unliganded BKV (A) and BKV-GT1b (B) structures. A single asymmetric unit of BKV (seen from outside the capsid) is shown coloured according to its local resolution in both cases. The interior of the asymmetric unit is also shown by rotating the asymmetric unit 180° about the y-axis. The highest resolution bin is 3.0 Å (blue) and the lowest resolution bin is 4.6 Å (red). A key is shown for reference. (Figure taken from (Hurdiss et al., 2018) [published under creative commons {CC BY} license]).

The resolution of the maps was sufficient to identify part of the shared C-terminus of VP2 and VP3, corresponding to residues 288-301 in the VP2 sequence (Figure 3-19D/E). Whilst all pentons contain symmetry-averaged density for VP2/3 (Hurdiss et al., 2016), only the pentavalent capsomeres contained density into which an atomic model could be built. This density is located at the base of the pentons as described previously for an isolated VP1 penton from PyV (X.S. Chen et al., 1998), and describes a small hydrophobic helix that is highly conserved in the VP2/3 sequences of polyomaviruses. The polypeptide sequence for this modelled region differs by four residues between BKV and PyV, one of which localises to the hydrophobic helix (X.S. Chen et al., 1998). Unlike PyV, there is much less interpretable density towards the N-terminus of this region for BKV, perhaps reflecting increased mobility in this solution structure of the intact virion.
Figure 3-18 – VP1 C-terminus mediated capsid assembly. EM density for residues 300-360 of the VP1 major capsid protein are shown at 2 $\sigma$ within 2 Å of the fitted coordinates. Chains 1-6 are coloured according to the scheme shown in Figure 3-9 and the five-fold, two-fold and three-fold symmetry axis are indicated by a pentamer, ellipse and triangle, respectively.
Figure 3-19 – Atomic model of the BKV capsid. (A) The T = 7d icosahedral model of the BKV capsid viewed down an icosahedral two-fold axis with each quasi-equivalent VP1 chain coloured according to the colour key shown. (B) A single asymmetric unit of the capsid showing the conformation of each of the C-termini. (C) Overlay of all the VP1 monomers in the asymmetric unit with the most variable regions (N-terminus, C-terminus, and CD loop) coloured according to chain number. (D) Internal view of a five-fold penton showing the model and EM density (2.2 σ within 1.6 Å of fitted coordinates) for the C-terminal region of VP2/VP3 binding at the base of the two adjacent VP1 monomers (grey). (E) Enlarged view of VP2/VP3 as shown in (D). (Figure taken from (Hurdiss et al., 2018) [published under creative commons {CC BY} license]).
3.4.3 Structural comparison to SV40

BKV and SV40 share 81.7% sequence identity in their major capsid protein, VP1, so as expected, the core fold of VP1 is very similar in the BKV structure to that previously seen in SV40. However, differences in the C-terminal arms of three of the six distinct conformers of VP1 were observed (Figure 3-20).

The biggest change is in chain 6 (grey in Figure 3-21A-B). In SV40, the electron density is only interpretable to residue Gln$^{355}$, where the polypeptide backbone points away from the virion (i.e. into bulk solution). In BKV, the polypeptide chain adopts a different conformation from Tyr$^{349}$ onwards, and in the EM density map this region is well resolved, allowing modelling of an additional four residues (up to Lys$^{359}$). This segment is resolved in an extended conformation that makes a series of previously unreported interactions with an arched structure formed by the C-terminus of chain 3 (green in Figure 3-21C). This is a significant difference with respect to the conformation seen for SV40, as this new conformation is consistent with the formation of up to 10 new H-bonding interactions.

Differences between BKV and SV40 are also observed at the icosahedral two-fold axes, which are formed at the interface between two adjacent chain 1s, in the region of residues 296-315 (Figure 3-22A). In BKV, the EM density is well resolved at the two-fold axis, allowing modelling of a new conformation of the chain 1 polypeptide that forms an intimate hydrophobic interface around which the two-fold axis of the particle is constructed (Figure 3-22B). This different conformation also results in a possible intramolecular interaction in BKV between Arg$^{214}$ (Lys in SV40) in the VP1 core and Asn$^{310}$ in the C-terminal arm (Figure 3-22C-D).
Figure 3-20 – Structural comparison of BKV and SV40 VP1. Each of the six quasi-equivalent VP1 chains from BKV are shown coloured according to the scheme used in Figure 3-9. The equivalent atomic coordinates from SV40 (black) are aligned to those of BKV. Pronounced differences in the N-terminus, C-terminus or CD loop between BKV and SV40 VP1 are indicated by asterisks.
Figure 3-21 – The C-terminus of BKV VP1. (A) Atomic model of the five-fold coordinated pentamer coloured as described in Figure 3-19. (B) Enlarged view of the C-terminus of chain 6 with the equivalent region from SV40 coloured black. (C) Enlarged view of the extended conformation shown in (B) showing potential hydrogen bonds between the C-terminus of chains 6 and 3 of BKV. (Figure taken from Hurdiss et al., 2018) [published under creative commons {CC BY} license].
Figure 3-22 – The two-fold symmetry axis of BKV. (A) Atomic model of two adjacent hexavalent pentamers which make up the icosahedral two-fold. (B) The symmetry-related C-termini of chain 1 indicated by a box in (A), showing the hydrophobic interaction between residues 296–304 with EM density for the sharpened map (3 σ). (C) A single VP1 monomer of chain 1 for BKV showing the equivalent C-terminus of SV40 (black). (D) Enlarged view of the possible intra-chain interaction denoted by a box in (C) showing EM density for the sharpened map low-pass filtered to 4 Å (2 σ). (Figure taken from (Hurdiss et al., 2018) [published under creative commons {CC BY} license]).
3.4.4 Location of disulphide bonds and Ca$^{2+}$ ions

Polyomavirus capsids are stabilised by Ca$^{2+}$ binding (J. Nilsson et al., 2005), and two putative locations for Ca$^{2+}$ binding have been previously identified by the presence of gadolinium ions in heavy metal soaking experiments for crystallographic structure determination of SV40 (Liddington et al., 1991). In the solution structures of BKV, a single density consistent with a Ca$^{2+}$ ion coordinated by glutamate sidechains and the backbone carbonyl of a serine residue can be seen (Figure 3-23A). This location is consistent with the proposed site 1 in SV40. However, no equivalent density for any metal ion binding at site 2 is present in any of the BKV structures (Figure 3-23A).

Disulphide bond formation is also critical for capsid stability (J. Nilsson et al., 2005), and their reduction/isomerisation in the endoplasmic reticulum of an infected host cell appears to be an essential step in virion uncoating (Schelhaas et al., 2007; M. Jiang et al., 2009). It has been shown previously that intermolecular disulphide bonds between Cys$^{104}$ residues serve to stabilise the SV40 capsid, and it has been suggested, but not shown, that Cys$^{9}$ may also be involved in disulphide bonding. The presence of a cysteine residue at position 104 is conserved between BKV and SV40, but the conformations of the CD loops which contain Cys$^{104}$ are different. In SV40, the CD loops of chains 2, 3 and 6 were observed in close proximity at an interface between pentavalent and hexavalent pentons, but that SV40 structure was not able to resolve which pair of chains participated in a disulphide bond (Stehle et al., 1996). In BKV, the conformation of the CD loops is different, and strongly suggests that a disulphide bond is formed between the CD loops of chains 3 (green) and 6 (grey). The CD loop of chain 2 (yellow) is orientated away from this interaction (Figure 3-23B). In contrast to SV40, in BKV the N-terminus of a chain 2 from an adjacent capsomere points towards the chain 2 CD loop. In the sharpened BKV maps, the first 13 residues of the chain 2 are not resolved, suggesting this portion of the capsid is less well ordered. However, in the unsharpened map, where high-resolution information is down-weighted and low-resolution features are more pronounced, a bridge of density joins the N-terminus to the CD loop before turning and continuing towards the encapsidated genome (Figure 3-23C). This density is entirely consistent with a previously undescribed Cys$^{104}$-Cys$^{9}$ disulphide bond. A similar bridge of density is observed beneath the icosahedral two-fold axes where
the N-terminus of chain 4 bridges to the CD loop of chain 1 (Figure 3-23E). As there is no high-resolution density for these interactions, the presence of Cys\textsuperscript{104} - Cys\textsuperscript{9} disulphide bonds was confirmed by determining the structure of BKV under reducing conditions (Figure 3-24A-C). The resulting structure is less ordered and has poorer resolution (6.5 Å) than the non-reduced structure (3.8 Å), likely reflecting the decreased rigidity of the virion. For direct comparison, the reduced and non-reduced structures were low-pass filtered to 7 Å. Crucially, the density for the Cys\textsuperscript{104}-Cys\textsuperscript{9} disulphides are completely absent in the reduced structure (Figure 3-23D & F).
Figure 3-23 – Location of calcium ions and disulphides in BKV. (A) The reported calcium binding sites 1 and 2 in the sharpened BKV map with EM density for calcium shown at 6 σ with surrounding residues coloured according to the corresponding chain. (B) Atomic model of the CD loops of chains 2, 3, and 6 and the N-terminus of an adjacent chain 2 from BKV with EM density shown at 1.5 σ. (C) Density from the unsharpened, non-reduced BKV map low-pass filtered to 7 Å (0.5 σ), showing additional density extending from the N-terminus of chain 2 and connecting to the CD loop of an adjacent chain 2. (D) Analogous region to that shown in (C) from the unsharpened reduced BKV map low-pass filtered to 7 Å (0.5 σ). (E) Density from the unsharpened BKV map low-pass filtered to 7 Å (1 σ), showing additional density extending from the N-terminus of chain 4 and connecting to the CD loop of chain 1. (F) Analogous region to that shown in (E) from the unsharpened reduced BKV map low-pass filtered to 7 Å (1 σ). (Figure taken from (Hurdiss et al., 2018) [published under creative commons {CC BY} license]).
Figure 3-24 – Reduced BKV data collection and cryo-EM map. (A) Typical micrograph from the reduced BKV data set (scale bar = 50 nm). (B) The plot of the Fourier shell correlation (FSC). Based on the 0.143 criterion for the ‘gold standard’ comparison of two independent data sets, the resolution of the reconstruction is 6.5 Å. (C) The 6.5 Å isosurface representation of reduced BKV viewed down an icosahedral two-fold axis (coloured according to the radial colouring scheme shown (Å)). (Figure taken from (Hurdiss et al., 2018) [published under creative commons {CC BY} license]).
The importance of the calcium ions and disulphide bonds identified in the BKV maps was further demonstrated by treating purified virions with reducing and chelating agents. Following incubation, negative stain EM examination revealed dissociated BKV capsids visible as dense clusters of capsomeres (Figure 3-25), presumably still bound to the viral genome by interactions mediated by the minor capsid proteins and the VP1 N-terminus described previously (Hurdiss et al., 2016). A small number of pentons were separated from these clusters, indicating they may be more loosely associated within the assembled capsid. A negative stain reconstruction of these revealed additional density protruding from the base of the penton, likely corresponding to VP2 or VP3 (Figure 3-26). Studies on SV40 have indicated that the five-fold coordinated pentons selectively dissociate during uncoating (Schelhaas et al., 2007). There is also evidence from the BKV data to support this mechanism of uncoating. During 3D classification of the BKV-GT1b data, a single class displayed weak density for the pentons at the capsid vertices, indicating partial occupancy or increased flexibility in this region (Figure 3-27). The same observation was observed for each of the BKV datasets reported in this section (data not shown).

Figure 3-25 – Dissociation of BKV under reducing and chelating conditions. (A) Negative-stain electron micrograph of purified BKV virions. B) Negative-stain electron micrograph of BKV virions following incubation with 5mM DTT and EDTA – scale bar 50 nm.
Taken together, these data suggest that despite having an apparently altered pattern of disulphide linkages within their capsids, both BKV and SV40 undergo selective dissociation of vertex pentons. Furthermore, negative stain reconstruction of dissociated pentons indicate that these remain associated with the minor capsid proteins. This suggests that interactions between the minor capsid proteins and the encapsidated genome may be weaker at the capsid vertices.

Figure 3-26 – Negative stain reconstruction of dissociated BKV pentons. (A) Negative-stain electron micrograph of BKV virions following incubation with 5mM DTT and 5mM EDTA – scale bar 50 nm. (B) Representative 2D class averages of fully dissociated pentons. (C) 3D reconstruction of the BKV penton with the atomic coordinates (PDB: 1CN3) of the PyV pentamer (blue) in complex with a fragment of VP2 (red) shown for scale. Additional density protruding from the penton base ascribed to the minor capsid proteins is indicated by an arrow.
Figure 3-27 – 3D classification of BKV. Following refinement of 40000 BKV-GT1b particles, the resulting particle stack was subject to 3D classification without particle alignment. The resulting four classes and their relative particle distribution are shown. An enlarged view of a five-fold coordinated pentamer is shown above each class, which revealed weak density in this region of the map for class three.
3.4.5 The interaction of BKV with GT1b oligosaccharide

GT1b is a branched oligosaccharide, described previously as having ‘left’ and ‘right’ arms ((Neu et al., 2013); Figure 3-28F). The disialic acid motif located on the ‘right’ arm, of GT1b is well resolved in the 3.4 Å map (Figure 3-28A-B). The binding mode is similar to that seen previously in a structure of GD3, an oligosaccharide that lacks a ‘left’ arm, bound to a BKV VP1 penton. The ‘left’ arm of GT1b is poorly resolved, indicating that it is not ordered at high-resolution. To examine the possible location of the GT1b ‘left’ arm, lower-resolution features in the map were emphasised by applying a 5 Å low-pass filter to the 3.4 Å map. This revealed bridges of density between Asp$^{59}$ and Lys$^{83}$ of VP1 and the GT1b density (Figure 3-28C), consistent with previous mutagenesis data of these residues which resulted in significantly reduced BKV spread in cell culture (Neu et al., 2013). In addition, molecular dynamics simulations were performed on a penton in complex with five GT1b molecules (Figure 3-30). For direct comparison with the EM density map the 3D density distribution of the GT1b heavy atoms was calculated from the accumulated 5.5 µs simulations (Figure 3-28D). A snapshot from this simulation reveals an orientation where Asp$^{59}$ and Lys$^{83}$ could interact with the ‘left’ arm of GT1b (Figure 3-28E). Additional interactions identified during the simulation are shown in Figure 3-29, and it is evident that the ‘left’ arm contributes to binding affinity through multiple transient interactions with the BKV surface (Figure 3-30).
Figure 3-28 – Interaction of BKV with GT1b. (A) Isosurface representation of a single five-fold pentamer of the BKV-GT1b complex with the capsid coloured as in Figure 3-19 and GT1b (magenta). (B) Enlarged view of the GT1b density (2 σ) containing the corresponding atomic model for the disialic acid motif of the right arm, coloured as in (D). (C) The 5 Å low-pass-filtered BKV-GT1b EM density (blue mesh) containing the high-resolution information shown in (B). (D) MD-derived atom density map for the simulated structure of GT1b overlaid with the 5 Å low-pass-filtered EM density shown in (C). (E) A snapshot of the MD simulated structure of GT1b showing possible interactions of the left arm with Asp and Lys. (F) SNFG-representation of GT1b oligosaccharide showing the left and right arms of the molecule (Varki et al., 2015). (Adapted from Hurdiss et al., 2018) [published under creative commons (CC BY) license]). Molecular dynamic simulations (D-E) performed by Dr Martin Frank, Biognos AB, Gothenberg. The sharpened and unsharpened BKV-GT1b cryo-EM map and corresponding atomic coordinates are available under the accession codes EMDB: 3944 and PDB: 6ESB, respectively.
Figure 3-29 – Molecular dynamics simulated interactions of BKV and GT1b. (A) VP1 residues which contact the left arm of GT1b during the simulation are coloured green, providing an interaction ‘footprint’ of the receptor molecule. The binding site for the well resolved disialic acid motif of the right arm is circled. (B) Bar graph showing the average number of contacts made by surface residues and GT1b during the simulation. Note: Neu5Ac (R) (Glycam residue label 0SA_5) was excluded from the analysis. (Figure adapted from (Hurdiss et al., 2018) [published under creative commons {CC BY} license]). Molecular dynamic simulations performed by Dr Martin Frank, Biognos AB, Gothenburg.
Figure 3-30 – Molecular dynamic simulations of GT1b. (A) A 3D-model of GT1b with Glycam residues labelled. (B) Graph showing the number of contacts made by the different components of GT1b with VP1 surface residues over time. Essentially, this shows the dynamics of the transient interactions made by the left arm of the GT1b molecule. (Figure adapted from (Hurdiss et al., 2018) [published under creative commons (CC BY) license]). Molecular dynamic simulations performed by Dr Martin Frank, Biognos AB, Gothenberg.
3.4.6 Interaction of BKV with glycosaminoglycans

To investigate recent reports of non-sialylated glycosaminoglycans acting as receptors for BKV infection, on-grid binding was performed with the same heparin used in recently published BKV inhibition assays (Figure 3-31A) (Geoghegan et al., 2017). The final sharpened map had a resolution of 3.6 Å (Figure 3-31B-C), and again there are no obvious structural differences in the BKV capsid compared to the BKV and BKV-GT1b structures. However, unlike the sharpened BKV-GT1b map, no obvious density for heparin was visible at high-resolution. To identify potential low-resolution modes of binding, difference maps were calculated between the low-pass filtered versions of the maps. The location of the resulting GT1b difference density is in agreement with the model of the branched GT1b molecule (Figure 3-32A), whilst density ascribed to heparin is located as a smear between the capsomers, and at the top of each capsomer pore (Figure 3-32B). The density between the capsomers locates the putative site of heparin binding atop a patch of positive charge on the virion surface (Figure 3-32D-E). At an increased threshold level, elongated density morphologically consistent with a linear heparin molecule remained visible above this positively charged patch, suggesting that this is where GAG molecules bind most tightly. This putative heparin density is located in the canyons surrounding the five-fold symmetry axis (Figure 3-33A) and in principle could accommodate eight saccharide units (Figure 3-33B). An atomic model of heparin was fitted as a rigid body within this density and positively charged residues within ~5 Å of the fitted coordinate were identified (Figure 3-34A). As BKV, JCV, SV40 and MCV have all been shown to bind GAG molecules (Geoghegan et al., 2017; Schowalter et al., 2011), the sequence conservation of these positively charged residues was compared across these species (Figure 3-34B). This demonstrated complete sequence conservation between BKV, JCV and SV40 for an RRXXR motif which forms a cluster of cationic residues surrounding the GAG density in BKV, and functional conservation of the remaining residues. On the other hand, less sequence and functional conservation was observed between BKV and MCV.
Figure 3-31 – BKV-heparin data collection and cryo-EM map. (A) Typical micrograph from the BKV-heparin data set (scale bar = 50 nm). (B) The plot of the Fourier shell correlation (FSC). Based on the 0.143 criterion for the ‘gold standard’ comparison of two independent data sets, the resolution of the reconstruction is 3.6 Å. (C) An isosurface representation of the 3.6 Å structure of BKV (coloured according to the radial colouring scheme shown (Å)). (Adapted from (Hurdiss et al., 2018) [published under creative commons {CC BY} license]).
Figure 3-32 – Interaction of BKV with heparin. (A) Isosurface representation of the unsharpened and 8 Å low-pass-filtered BKV-GT1b map viewed down the icosahedral two-fold axis and coloured according to the radial colouring scheme shown (Å). The GT1b difference density (3.4 σ) is shown coloured magenta. (B) The unsharpened and 8 Å low-pass-filtered isosurface representation of the BKV-heparin structure, with the putative heparin difference density (3.4 σ) coloured orange. (C) Enlarged section of the BKV capsid with the surface coloured according to electrostatic surface potential (N.A. Baker et al., 2001). (D) The same view shown in (C) overlaid with the putative heparin difference density shown in (B) (orange). The positively charged intercapsomere region is circled and the fivefold symmetry axes are denoted with black pentamers. (Adapted from (Hurdiss et al., 2018) [published under creative commons {CC BY} license]).
Figure 3-33 – Putative heparin density surrounding the capsid vertices. (A) 32 Å thick slab through the unsharpened and 8 Å low-pass-filtered BKV-heparin structure containing the fitted atomic coordinates (grey), viewed down the icosahedral five-fold. Elongated difference density consistent with heparin (orange) is shown at 3.7 σ and contains the fitted coordinates of eight saccharide units of heparin, generated from PDB: 3IRJ. (B) Enlarged view of a single fitted heparin molecule.
Figure 3-34 – Sequence conservation of proposed heparin binding residues. (A) Surface representation of the putative heparin binding site on BKV containing the corresponding atomic model of the capsid shown in ribbon format (grey). Difference density ascribed to heparin is shown as an orange mesh (3.8 σ) and contains the fitted atomic model of eight saccharide units of the GAG molecule (black). Positively charged residues within 5 Å of the fitted heparin molecule are shown and labelled. (B) Table displaying sequence conservation between BKV residues shown in (A) and those of related viruses (JCV, SV40 and MCV). Green indicates complete sequence conservation, blue - functional conservation to another positively-charged residue and red – no conservation.
3.5 Discussion part two

The structures of BKV presented here provide a detailed view of a native, infectious, human polyomavirus. The architecture of the BKV capsid is essentially the same as previously described for SV40 and PyV. However, the quality and completeness of the cryo-EM structures means that different conformations of the C-termini of VP1 can be described for a particle in solution. The C-termini play a fundamental role in the assembly and stability of a polyomavirus capsid, adopting six distinct conformations that allow a single VP1 sequence to adopt all of the positions in a $T=7d$ lattice. It is unclear whether these differences arise from sequence differences between SV40 and BKV or from the different solution conditions of the experiments. It should be noted however, that the solution structures determined herein, are in more native-like conditions than those used to determine the crystal structures (Lattman, 1980; Stehle et al., 1996). So, whether differences arise from experimental conditions or sequence variation, they provide new insights into the structure of a human pathogen in near-native conditions. The new inter-subunit interactions described here Figure 3-22 have implications for the structure and stability of the virus. The conformation observed for the C-terminus of Chain 6 (Figure 3-22B) suggests that it interacts with a loop of the C-terminal arm of chain 3, generating up to 10 potential H-bonding interactions (Figure 3-22C). Given that this interaction occurs sixty times in an intact capsid, this represents a significant source of additional stability for the BKV virion. In addition, amino acid variations between BKV genotypes map to this region of the C-terminus of VP1, indicating that they may play contributing role in immunogenicity (Pastrana et al., 2013).

A different conformation is also evident in the C-termini of chain 1, which make the two-fold axes of the BKV and SV40 capsids (Figure 3-22D-G), forming a tighter and more intimate hydrophobic interface. Interestingly, additional differences are also seen beneath (at lower radius) this interface. The N-terminus of chain 4 and the CD loop of chain 1 appear to make a previously unreported disulphide bond (Cys$^9$-Cys$^{104}$), which would further stabilise the capsid in this region. These structures refine the model of how pentavalent and hexavalent pentons are linked by disulphide bonds, by showing that only disulphides between Cys$^{104}$ residues in chain 3 and chain 6 are present. The previous
suggestion that chain 6 might also disulphide bond to chain 2 would not be possible in the BKV structures. Furthermore, the chain 2 CD loop is in a different conformation to that seen in SV40, that allows disulphide bonding to the N-terminus of a second chain 2 molecule. These observations account for the disulphide bonding of all of the CD loops in the structure, and two of the N-termini. Therefore, a much more complete picture of disulphide bonding in polyomavirus capsids is reported. These differences could reflect genuine differences between BKV and SV40 or indicate that the SV40 in the crystals from which the structure was solved was not in a fully oxidised state.

Structural data is also presented for two different receptors bound to the surface of the BKV. In the 3-4 Å resolution range of the structures presented, there is no detectable conformational change in VP1. In the BKV-GT1b structure, the disialic acid motif of the GT1b ‘right’ arm is resolved to high-resolution, confirming a conserved mode of binding between the large B-series gangliosides and smaller variants studied previously. MD simulations show that the ‘left’ arm of GT1b is dynamic, explaining the lack of high-resolution density in the EM map, and its previous intractability to crystallographic studies. Low-resolution EM density and MD suggest a likely mechanism of binding for the ‘left’ arm of GT1b, in which multiple, weak interactions occur with the VP1 surface. Such interactions provide the structural correlate to previous mutagenic and functional studies showing that the presence of a ‘left’ arm enhances infectivity (Neu et al., 2013).

These results demonstrate the emerging role cryo-EM can play in structural glycobiology, to complement existing NMR and crystallographic approaches. The ability to produce polyomavirus-ligand structures at resolutions approaching that of X-ray crystallography also leads to the exciting prospect that cryo-EM could be used to study small molecule inhibitors of BKV infection or assembly, and indeed the structure of BKV-GT1b presented here is an ideal platform for the in silico design of such tool molecules. It is clear that neither GT1b nor GAG binding induce detectable conformational change in the VP1 component of the capsid, and presumably act simply as attachment receptors to facilitate cell entry, where changes in redox state and pH can effect changes in capsid structure that lead to uncoating (M. Jiang et al., 2009).

It is not possible to determine whether the density between pentamers represents bound GAG molecules, but the lack of resolution suggests that no single mode
of binding exists. However, this site displays a positive charge, as do GAG binding sites observed in other viruses (O'Donnell et al., 2009; Fry et al., 1999). Furthermore, this inter-capsomere density is morphologically consistent with a long unbranched polysaccharide molecule of approximately eight units in length. Also, the charged residues surrounding the proposed heparin density are highly conserved between BKV, JCV and SV40, which have all been found to use GAGs for attachment (Geoghegan et al., 2017). The lack of conservation between BKV and MCV suggests a different mode of heparin binding for this more distantly related virus, which may rationalise its increased dependency on GAGs (Schowalter et al., 2011). From an immunological perspective, GAG binding in the ‘canyon’ between pentamers would make sense as this would reduce accessibility to neutralising antibodies. Density in the pore of each pentamer and at the two-fold axes would be subject to inappropriate symmetry averaging, and so must be interpreted with great care. The possibility that GAGs bind in the pore of the VP1 pentamer is intriguing however. There is no clear evidence for an interaction between GAGs and VP2/3, but the pore is directly above their location, and it is interesting to note that BKV pseudo-viruses that lack VP2 and VP3 are less efficient at transducing a range of different cell types (Schowalter and Buck, 2013), which might be partially explained by a reduced affinity for GAGs. Indeed, the entire VP1 shell is remarkably porous, potentially allowing other opportunities for the interactions between minor capsid proteins and GAGs that have been recently reported for the closely related papillomaviruses (J. Guan et al., 2017). These observations provide structural clues about GAG binding to BKV and may form a basis to determine the precise molecular mechanism of GAG interaction using shorter fragments of heparin, which may be more amenable to high-resolution structural characterisation. Based on the preliminarily observations described here, a GAG sample with a minimum length of eight saccharides would most likely be required. Furthermore, mutational analysis of the conserved cationic residues identified here may help elucidate which of these are required for GAG binding. GAG analogues have reportedly been used to treat BKV-associated disease (Van der Aa et al., 2014; Winter et al., 2015; Isik et al., 2014; Cervigni, 2015). The rationale for such treatments was based on restoration of the barrier function of the bladder epithelium, but in light of recent results (including those presented here), it is possible that GAGs may bind directly to BKV and perturb cellular attachment or entry.
The rationale for the studies presented here is to inform the design of future anti-BKV therapies. Such therapies could include antibodies capable of neutralising BKV, or BKV vaccination. The first human mAbs which bind BKV virions have recently been reported (Jelcic et al., 2015), and vaccination with BKV VLPs has been shown to induce pan-specific immunity in mouse models (Pastrana et al., 2013). The structures of the native, reduced, and receptor-bound virions provide a platform for understanding such antibody responses to BKV, and features that underpin vaccine stability. Going forward, co-structures with mAbs that show BKV serotype-specific neutralisation (Randhawa et al., 2009), broad cross-neutralisation, or non-neutralising mAbs will be key for guiding efforts to further develop therapies to protect patients against BKV nephropathy (Buck, 2016).
Chapter 4 Structural studies of Coxsackievirus A24v

4.1 Introduction

The Enterovirus genus, within the Picornaviridae family, comprises 13 species, seven of which contain pathogens of humans (Figure 4-1). Poliovirus is undoubtedly the most well-known and characterised enterovirus, having been the focus of decades of research (Racaniello and Baltimore, 1981a; Racaniello and Baltimore, 1981b; Hogle et al., 1985). Excluding the three poliovirus serotypes, there over 250 enteroviruses within the seven pathogenic species. These include rhinoviruses, numbered enteroviruses, echoviruses and coxsackieviruses (Baggen, Thibaut, et al., 2018).

<table>
<thead>
<tr>
<th>Genus</th>
<th>Species</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterovirus</td>
<td>Enterovirus A</td>
<td>Enterovirus A71</td>
</tr>
<tr>
<td></td>
<td>Enterovirus B</td>
<td>Coxsackievirus B3</td>
</tr>
<tr>
<td></td>
<td>Enterovirus C</td>
<td>Coxsackievirus A24v</td>
</tr>
<tr>
<td></td>
<td>Enterovirus D</td>
<td>Enterovirus D68</td>
</tr>
<tr>
<td></td>
<td>Rhinovirus A</td>
<td>Rhinovirus A2</td>
</tr>
<tr>
<td></td>
<td>Rhinovirus B</td>
<td>Rhinovirus B14</td>
</tr>
<tr>
<td></td>
<td>Rhinovirus C</td>
<td>Rhinovirus C15</td>
</tr>
</tbody>
</table>

Figure 4-1 – Enteroviruses of humans. The seven species of enteroviruses which contain human pathogens, are depicted, with an example of a clinically significant genotype shown next to each. (Figure adapted from (van der Linden, Wolthers, et al., 2015) [published under creative commons {CC BY 4.0} license].)

The respiratory and faecal-oral route are the primary modes of transmission for enteroviruses (Pons-Salort et al., 2015). Enterovirus may spread from their primary sites of replication in the gastrointestinal or respiratory tract and infect other tissues and organs such as the central nervous system (Bruu, 2002). Enterovirus infection may be asymptomatic but can also cause a myriad of diseases, with wide-ranging severity. Such infections most often affect
immunocompromised individuals, infants and young children. Enteroviruses are not typically responsible for gastrointestinal diseases, as their name would suggest. The numbered enteroviruses, echoviruses and Coxsackieviruses are the primary cause of aseptic meningitis. In addition, they are responsible for diseases such as hand-foot and-mouth disease (HFMD), acute flaccid paralysis (AFP), neonatal sepsis-like disease, non-specific febrile illness, herpangina, pleurodynia, myocarditis and pericarditis (Pons-Salort et al., 2015). As well as being the causative agent of the common cold, rhinoviruses also cause exacerbation of chronic obstructive pulmonary disease (COPD) and asthma, in addition to respiratory tract disease (Thibaut et al., 2016). Furthermore, echoviruses and Coxsackieviruses B also contribute to the aetiology of type 1 diabetes as a result of the inflammation caused by their persistent infection of pancreatic β-cells (Hober and Sauter, 2010).

The enteroviruses EV-D68 and EV-A71, have emerged as significant risks to public health, being responsible for disease outbreaks worldwide. EV-D68 replicates in the respiratory tract and was historically considered a rare pathogen but is now increasingly been detected in patients with severe respiratory disease. In 2014, EV-D68 was responsible for a nationwide outbreak of severe respiratory disease in the United States which also saw infected individuals suffering from AFP and cranial nerve dysfunction (Midgley et al., 2015; Messacar et al., 2015; Holm-Hansen et al., 2016). A causative agent of HFMD, EV-A71, has been responsible for huge outbreaks across Southeast Asia. Whilst most infections are not severe, life threatening complications may include meningitis and brainstem encephalitis (Yi et al., 2017; S.-M. Wang and Ching-Chuan Liu, 2014).

Another well-known pathogenic enterovirus, and the focus of work described in this chapter, is CV-A24v which belongs to the enterovirus C species. CV-A24v is the main aetiological agent of AHC, and is responsible for >10 million AHC cases worldwide (Oh et al., 2003; Tavares et al., 2006). AHC is characterised by a sudden onset of ocular pain, swelling, watering, and extensive subconjunctival haemorrhaging. Since AHC is highly contagious it spreads rapidly and can affect up to 48% of a population in a single outbreak (Ghazali et al., 2003). Coxsackievirus A24 (CV-A24) was first isolated in South Africa in 1951 (Measroch and Gear, 1951) but at this point it was not associated with disease in humans. However, in 1970, a pathogenic variant of CV-A24 emerged that was
responsible for a large AHC outbreak in Singapore (Kono et al., 1972). Initially, this AHC-causing variant circulated only in Southeast Asia but subsequently spread worldwide, suddenly causing numerous explosive AHC outbreaks and two pandemics, (starting in 1985 and 2002 (Miyamura et al., 1988; Lévêque et al., 2010; Ishiko et al., 1992)). Although both the variant and nonvariant strains of CV-A24 are being detected in clinical samples worldwide, only strains belonging to the “variant” clade are associated with AHC. As of yet, the molecular mechanism underlying this difference in pathogenicity is unknown. Also, it remains unclear what has led to the sudden onset of CV-A24v pandemics.

There are over 100 serotypes of non-polio enteroviruses alone and as such, vaccination to all of these is not a viable option (Baggen, Thibaut, et al., 2018). Thus, there is an urgent clinical need for the development of broad spectrum, anti-enterovirus drugs (Thibaut et al., 2012). Such efforts will be bolstered by structural and functional studies of pathogenic enteroviruses, such as CV-A24v, which provide the molecular details of viral targets for inhibition.

4.1.1 Genome organisation, translation and processing

Enteroviruses are small (~30 nm in diameter), non-enveloped viruses with a positive sense single-stranded RNA genome of approximately 7.5 kb in length. The genome encodes a polyprotein which encodes four structural proteins (VP1-4) in the P1 region and seven non-structural proteins (2A-C, 3A-D) in the P2 and P3 regions (Figure 4-2). The genome comprises a single open reading frame with a 5’ untranslated region, which is highly structured, and a polyadenylated 3’ end. The viral protein 3B is covalently linked to the 5’ end of the viral genome and is usually referred to as VPg (viral protein genome-linked). An internal ribosome entry site (IRES) is located at the 5’ end of enterovirus genome which facilitates cap-independent translation. Following translation, the 2A, 3C, and 3CD proteases proteolytically process the polyprotein into its individual protein components (van der Linden, Wolthers, et al., 2015; Baggen, Thibaut, et al., 2018).
4.1.2 Enterovirus structure

The first animal enteroviruses to be characterised structurally were poliovirus and rhinovirus-14 (Hogle et al., 1985; Rossmann et al., 1985). Since these 3D structures were elucidated, a number of enterovirus have been structurally characterised which revealed a common capsid architecture. Enteroviruses have a pseudo $T=3$ icosahedral capsid, the outside of which is formed by the structural proteins VP1, VP2 and VP3 (Figure 4-3). These proteins have a typical ‘jelly-roll’ fold formed from opposing, four-stranded antiparallel β-sheets which create a β-barrel with a hydrophobic core. The viral genome and the small, myristoylated VP4 protein are located inside of the particle (Zocher et al., 2014; Plevka et al., 2012; Hendry et al., 1999). The high antigenic diversity of enteroviruses is largely dictated by highly variable loops on the capsid surface which are readily accessible to host neutralising antibodies. For the majority of enteroviruses, the five-fold symmetry axis is surrounded by a surface depression known as the canyon which is a known site for receptor binding (Rossmann et al., 2002). For most enteroviruses, a hydrophobic cavity formed by VP1 at the base of these canyons contains a lipid molecule known as the ‘pocket factor’ (Smyth et al., 2003; Smith et al., 1986; Filman et al., 1989). The purpose of these fatty acid molecules is thought to be particle stabilisation during transmission and as a cue
for uncoating (X. Wang et al., 2012). These molecules usually comprise a linear lipid of varying length (Smyth et al., 2003), however CV-A24v was shown to have an atypical, branched pocket factor (Zocher et al., 2014) and rhinovirus C lacks one entirely (Yue Liu et al., 2016). Whilst most enterovirus capsids have a similar capsid architecture, some have unique surface features. For example, in contrast to the generally smooth and featureless appearance of most enteroviruses, the recently elucidated rhinovirus C has a spiked appearance which arises from “finger” like protrusions at the juncture between VP1, VP2 and VP3. Furthermore, both rhinovirus C and enterovirus D68 have canyons which are narrower than those observed in other enterovirus but the functional significance of this is yet to be determined (Yue Liu, Sheng, Fokine, et al., 2015; Yue Liu et al., 2016).

![Figure 4-3 – Enterovirus structure. Enterovirus capsids comprise sixty copies of the four different structural proteins (VP1-VP4) which are arranged as a pseudo T=3 icosahedral particle (Figure adapted from Yi et al., 2017) [published under creative commons (CC BY-NC 4.0) license].](image)

### 4.1.3 The enterovirus life cycle

#### 4.1.3.1 Receptor binding

Enteroviruses recognise a broad range of receptors, which likely partially accounts for the wide range of different diseases they cause. Enterovirus receptors can be divided into two categories which are attachment and uncoating.
Attachments receptors, as their name suggests, promote attachment of viruses to the host cell surface and are believed to facilitate virus uptake. Sialylated glycans and glycosaminoglycans are known attachments receptors for several enteroviruses (Baggen, Thibaut, et al., 2018). For example, CV-A24v has been shown to bind to an α2,6-linked sialic acid molecule (Figure 4-4), with the binding site situated close to the five-fold symmetry axis (E.C. Nilsson et al., 2008; Zocher et al., 2014). Enterovirus D70 (EV-D70) also utilises sialylated glycans for cellular attachment, but in contrast to CV-A24v, it recognises an α2,3-linked sialic acid molecule. Furthermore, EV-D70 has large sequence differences in the receptor binding region identified in CV-A24v and is therefore unlikely to recognise sialic acid in the same way (Nokhbeh et al., 2005). Proteins can also serve as attachment receptors for enteroviruses. For example, a number of enterovirus and echoviruses have been found to interact with the glycosylphosphatidylinositol (GPI)-anchored surface protein, decay-accelerating factor (DAF) (Bergelson et al., 1994). Binding of DAF by group B coxsackieviruses was shown to enhance infection by liberating the uncoating receptor, coxsackievirus–adenovirus receptor (CAR), from tight junctions through a series of signalling events (Coyne and Bergelson, 2006). Furthermore, some rhinoviruses can utilise the low-density lipoprotein receptor (LDLR) as an attachment receptor (Hofer et al., 1994), which binds near to the icosahedral five-fold, similarly to the sialic acid binding site observed in CV-A24v (Zocher et al., 2014).
Uncoating receptors are typically proteins belonging to the immunoglobulin-like or integrin receptor family (Rossmann et al., 2002). These molecules bind to the enterovirus canyon region (Figure 4-5) and cause destabilisation of the particle, resulting in conformation changes in the capsid which ultimately lead to genome uncoating (X. Wang et al., 2012). The first uncoating receptors to be discovered were the poliovirus receptor (PVR) (Mendelsohn et al., 1989), and intercellular adhesion molecule-1 (ICAM-1), used by the major group rhinoviruses (Greve et al., 1989) and Coxsackievirus A21 (CV-A21) (Lonberg-Holm et al., 1976; Shafren et al., 1997). The causative agents in HFMD, EV-A71 and Coxsackievirus A16 (CV-A16), use the scavenger receptor class B member 2 (SCARB2) as an uncoating receptor (Yamayoshi et al., 2009; Yamayoshi et al., 2012). Given that SCARB2 is abundantly expressed on glial cells and neurons, this may explain the neurological diseases associated with EV-A71 (Jiao et al., 2014; Fujii et al., 2013). Whilst most uncoating receptors are proteins, EV-D68 has been shown to interact with sialic acid, which bind to the canyon and induce uncoating (Yue Liu, Sheng, Baggen, et al., 2015; Baggen et al., 2016). The interaction between many important human enteroviruses and their uncoating receptors have yet to be structurally characterised (e.g. EV-A71 and SCARB2). In addition, many of the structures which are available are low-resolution cryo-EM maps, such as CV-A21.
and rhinoviruses in complex with ICAM-1 which range from ~8 to 20 Å resolution (Xiao et al., 2001; Xiao et al., 2005; Kolatkar et al., 1999; Xing et al., 2003).

**Figure 4-5 – Uncoating receptor interactions with the canyon.** (A) Coxsackievirus–adenovirus receptor (purple) bound to the canyon of Coxsackievirus B3 – PDB 1JEW (He et al., 2001). (B) Intercellular adhesion molecule-1 (purple) bound to the canyon of Coxsackievirus A21 - PDB 1Z7Z (Xiao et al., 2005). (Figure adapted from (Zocher et al., 2014) [published under creative commons {CC BY 4.0} license]).

### 4.1.3.2 Uncoating

Enterovirus uncoating results from progressive destabilisation of the capsid in response to cellular cues, which leads to genome release during cellular entry (Figure 4-6). There are three distinct types of enterovirus particle which were originally identified by studies on poliovirus. These different particles have different antigenic properties and are distinguished by their different sedimentation coefficients. They are typically named 160S, 135S and 80S (Belnap et al., 2000). The native, infectious form of the virus which contains all four structural proteins is called 160S. It is thought that the 160S particle undergoes a process of “breathing” where dynamic shrinkage and expansion of the capsid results in transient exposure of the VP1 N-terminus and VP4 (Q. Li et al., 1994). The 135S particle, also known as the altered particle, or A-particle, is formed in response to a variety of cellular cues, depending on the particular enterovirus in question. Such triggers of uncoating include receptor binding and low-pH, such as that encountered in the endosome. The 135S particle is expanded approximately 4% compared to the native particle and has lost its internal protein VP4 (Belnap et al., 2000). The N-terminus of VP1, which contains a predicted amphipathic helix, is also exposed, rendering the surface of the
enterovirus particle hydrophobic, allowing it to interact with membranes (Fricks and Hogle, 1990). These structural rearrangements are also associated with a redistribution of the genome within the capsid, likely in preparation for release (Ren et al., 2013; Lyu et al., 2014). During cellular entry, the first uncoating cues which enteroviruses encounter are receptors. These molecules interact with the VP1 GH loop, situated above the canyon, and induce a conformational change, resulting in collapse of the hydrophobic pocket and displacement of the stabilising pocket factor (X. Wang et al., 2012). Some enteroviruses, such as the minor group rhinoviruses, rely on the late endosomal low pH as a cue for uncoating. The use of this chemical trigger presumably achieves spatial and temporal regulation during the uncoating process (Prchla et al., 1994). Studies on EV-A71 have demonstrated a dual dependency on receptor binding and low pH to trigger uncoating (Hussain et al., 2011). Upon genome release, the non-infectious 80S particle is formed (Baggen, Thibaut, et al., 2018). Whilst 135S and 80S particles are available for a number of enteroviruses (e.g. EV-A71, CV-A16, CV-A6 and CV-B3), these remain to be characterised for CV-A24v (Xu et al., 2017; Lyu et al., 2014; Ren et al., 2013; H. Lee et al., 2016).
Figure 4-6 – Enterovirus uncoating stages. Upon encountering an uncoating stimulus, the mature enterovirus undergoes conformational change which results in the release of VP4 and exposure of the VP1 N-terminus. Another trigger causes the genome to be released, leaving behind the empty capsid (Figure adapted from (Shingler et al., 2013) [published under creative commons {CC BY 4.0} license]).
4.1.3.3 Pore formation

Enteroviruses are non-enveloped and as such cannot utilise membrane fusion machinery to deliver their genome into the cytosol. Whilst there is a wealth of information on the structural changes which occur in enterovirus particles during uncoating, less is known about the process of genome translocation. It has been demonstrated that recombinant, myristoylated VP4 from rhinovirus 16 oligomerises and forms pores in membranes, the diameter of which are sufficient to allow passage of a single stranded RNA molecule (Davis et al., 2008; Panjwani et al., 2014). Studies have shown that poliovirus particles can interact with lipid membranes to form ion-permeable channels, and that mutations in VP4 can interfere with this process (Tosteson and Chow, 1997; Danthi et al., 2003). More recent electron tomography experiments suggest that poliovirus particles are connected to receptor decorated liposomes by a 50-Å long connector known as the ‘umbilicus’ (Strauss et al., 2013). It is thought that this structure serves the purpose of protecting viral RNA during translocation. This hypothesis is supported by recent studies which demonstrated that the genome of poliovirus is protected from co-endocytosed RNase enzymes (Groppelli et al., 2017).

4.1.3.4 Genome release

It was originally thought that the five-fold symmetry axis of the enterovirus capsid was the site of genome egress (Hadfield et al., 1997). Subsequently, structures of 135S and 80S particles were determined by X-ray crystallography, which revealed an opening which extends across the two-fold symmetry axis and the quasi-three-fold, where the VP1 N-terminus is externalised (X. Wang et al., 2012; Ren et al., 2013; Lyu et al., 2014). In addition, electron microscopy studies of heated poliovirus and rhinovirus particles indicate that the two-fold symmetry axis is the probable site of genome egress (Bostina et al., 2011; Harutyunyan et al., 2013). Most in vitro studies on enterovirus genome egress have been carried out under conditions which globally stimulate particles (e.g. with excess of receptor in solution or low pH conditions) (Xing et al., 2003). In a true infection, the process of genome egress is likely to be asymmetric as only one side of the enterovirus particle is likely to interact with cell surface receptors and endosomal membrane. A recent study began to address the question of asymmetric particle stimulation by determining the structure of CV-B3 in complex with a membrane embedded receptor molecule (H. Lee et al., 2016). In contrast to globally stimulated particles,
this asymmetric structure revealed an opening at the three-fold symmetry axis of the capsid. Future studies into the process of enterovirus uncoating will likely require examining this process within infected cells. Such experiments are now possible as a result of the advancements in the field of in situ cryo-electron tomography (cryo-ET), but they remain a prodigious undertaking. Another lingering question in the entry process of enteroviruses is the directionality of genome egress. Currently, there is one study on this process, which demonstrated that the 3’ end of rhinovirus-A2 exits the capsid before the 5’ end (Harutyunyan et al., 2013). There is further speculation about the role of additional factors which may regulate the release of the viral genome. One such candidate is the host factor phospholipase A2 group XVI (PLA2G16). The physiological role of PLA2G16 has not been fully elucidated, however, it was shown to be a pan-enterovirus uncoating factor that mediates the release viral genomes from within endocytic vesicles (Staring et al., 2017).

4.1.3.5 Genome replication

Enterovirus replication is a highly conserved process (Paul and Wimmer, 2015). The non-structural protein 3D, an RNA-dependent RNA polymerase, is the key enzyme in enterovirus replication. Higher-order RNA structures located on the 5’ and 3’ ends of the genome, and the viral protein 3B (VPg), form the site of replication initiation. These priming elements are bound by 3D and the non-structural proteins 3CD and 3AB. In addition, 3D interacts with the multifunctional viral protein 2C, which functions as an ATP-dependent and also ATP-independent RNA chaperone (Xia et al., 2015). A number of host proteins are also involved in enterovirus replication, including heterogeneous nuclear ribonucleoprotein C, poly(rC)-binding protein 2 and polyadenylate-binding protein 1 (Baggen, Thibaut, et al., 2018). Collectively, these viral and host proteins are known as replication complexes (RCs). The higher-order organisation of these RCs remains enigmatic but it seems likely that techniques such as cryo-EM and cryo-ET will help elucidate this in the near future.

The non-structural proteins 2B, 2C and 3A cooperatively hijack host cell proteins and alter host cell membranes and lipid homeostasis to generate membranous replication organelles (ROs), which serve as platforms for genome replication and virion morphogenesis (Suhy et al., 2000). The lipid kinase, phosphatidylinositol 4-kinase-β (PI4KB), is recruited to ROs by the 3A protein via an interaction with
acyl-CoA-binding domain-containing protein 3. PI4KB recruitment is also proposed to involved the viral protein 2BC (Arita, 2014). This results in an increase in phosphatidylinositol-4-phosphate (PtdIns4P) lipids within the ROs, which in turn leads to recruitment of oxysterol-binding protein and a subsequent build-up of cholesterol (van der Schaar et al., 2016; Strating and van Kuppeveld, 2017). The accumulation of cholesterol and PtdIns4P is proposed to facilitate RO formation, correct processing of the viral polyprotein and efficient genome replication (Ilnytska et al., 2013; Ford Siltz et al., 2014; Arita, 2016; Melia et al., 2017). Whilst the role of 3A in RO formation has been well characterised, the same is not true of 2B and 2C. It has been shown that 2C binds to reticulon 3, a host protein implicated in membrane curvature, but this interaction and its significance has not been examined (W.-F. Tang et al., 2007). It is known that 2B functions as a viroporin, but how this relates to the formation of ROs is not certain (Agirre et al., 2002; van Kuppeveld et al., 2005).

Figure 4-7 – Enteroviruses remodelling of cellular membranes. (A) A slice through a serial tomogram collected on a Coxsackievirus B3 infected cell (5 hours post infection). (B) Surface rendering of the section indicated by a white box in (A), showing the surface topography of virally induces membranous structures (green, yellow and orange), adjacent to the endoplasmic reticulum (blue) (Limpens et al., 2011). Scale bar = 500 nm. (Figure taken from (van der Linden, Wolthers, et al., 2015) [published under creative commons {CC BY 4.0} license]).
4.1.3.6 Virion assembly

The process of capsid morphogenesis remains poorly understood for enteroviruses (P. Jiang et al., 2014). Initially, the myristoylated capsid precursor protein P1 is proteolytically cleaved into VP0, VP1 and VP3 by 3CD protease, a process which is catalysed by the cellular chaperone heat shock protein 90 (HSP90) (Geller et al., 2007). These cleaved components self-assemble into the viral asymmetric unit, five of which oligomerise to form pentamers. Studies have shown that myristate molecules (Paul et al., 1987; Chow et al., 1987) and, in a number of cases, glutathione stabilise these pentamers (Thibaut et al., 2014; H.-C. Ma et al., 2014). Broadly, there are two general models of enterovirus assembly (Figure 4-8). The first suggests that procapsids, comprising VP1, VP2 and VP0, are formed from twelve pentamers and then the viral genome is inserted through some currently unknown mechanism (Phillips et al., 1968; Oppermann and Koch, 1973). The second model proposes that provirions are formed when pentamers assemble around viral RNA (Ghendon et al., 1972). In the case of the second hypothesis, procapsids are merely an off pathway assembly by-product which serve as a reservoir of capsid components in infected cells (Shingler et al., 2013). Regardless of their role within enterovirus morphogenesis, procapsids are readily isolated from enterovirus infected cells (Curry et al., 1997; Basavappa et al., 1994; Chia-Chyi Liu et al., 2011). The capsid protein VP3 is known to interact with the 2C protein, a crucial component of the viral replication complex. This protein-protein interaction is believed to govern the specificity of encapsidation (Ying Liu et al., 2010). An alternative theory is that the viral genome contains RNA packaging signals which preferentially bind to the viral coat proteins, thus directing capsid assembly. Whilst packing signals have not yet been identified in enterovirus genomes, these have been identified in the related parechovirus 1 (Shakeel et al., 2017) and FMDV (Logan et al., 2017). Such observations warrant further investigation into the possible role of RNA packaging signals in enterovirus assembly. Finally, the mature enterovirus particle is formed following the RNA mediated cleavage of VP0 to form VP2 and VP4 (Basavappa et al., 1994).
Figure 4-8 – Proposed models of enterovirus assembly. In model one, twelve pentamers assemble to form an empty procapsid (1A) and the genome is then inserted (1B). In the second model, a provirion is formed when twelve pentamers assemble around the viral genome (2). In both cases the final maturation step is the autocatalytic cleavage of VP0 to form VP2 and VP4 (Figure adapted from Shingler et al., 2013 [published under creative commons (CC BY 4.0) license]).

4.1.3.7 Virion release

Newly assembled enterovirus particles are generally thought to be released by lytic mechanisms (Figure 4-9). However, there is an emerging body of evidence which suggests that enteroviruses can also be released by non-lytic means (Y.-H. Chen et al., 2015; Too et al., 2016). A number of studies have highlighted the role of autophagy in enterovirus release (Lai et al., 2016; Robinson et al., 2014; Bird et al., 2014). Double-membrane autophagosomes are able to capture enterovirus particles and subsequently fuse with the plasma membrane of the cell to release single-membrane vesicles (Jackson et al., 2005). These extracellular vesicles are rich in phosphatidylserine which facilitates their cellular uptake by phosphatidylserine receptors. However, given that the enterovirus receptor is still required for infection, it is possible that virions are released prior to cellular entry (Y.-H. Chen et al., 2015; Baggen, Thibaut, et al., 2018). This process of non-lytic release of enteroviruses is reminiscent of Hepatitis A virus (HAV) which exists in both an enveloped and non-enveloped form (McKnight et
The enveloped form of HAV is resistant to neutralising antibodies and the same may hold true for enteroviruses.

**Figure 4-9 – Enterovirus replication cycle.** Schematic representation of the enterovirus lifecycle (Figure adapted from (van der Linden, Wolthers, et al., 2015) [published under creative commons (CC BY 4.0) license]).

### 4.1.4 Epidemiology of CV-A24v

AHC is characterised by the quick onset of an extremely painful conjunctivitis and subconjunctival haemorrhaging. Large outbreaks of AHC are chiefly caused by CV-A24v and EV-D70. Direct or indirect eye contact with eye secretions is believed to be the primary mode of spread between individuals. However, it has also been suggested that CV-A24v may be transmitted through the respiratory route (Yin-Murphy et al., 1993), and the presence of these strains in stool samples indicates that the faecal-oral route may also play a role in virus spread (Fonseca et al., 2012; Kono et al., 1977; E.C. Nilsson et al., 2008). Cases of AHC typically resolve within a week, but complications from microbial superinfection can occur. In addition to AHC, these viruses can cause symptoms in the upper respiratory tract, cornea and nervous system, such as acute flaccid paralysis in
a small number of cases (Kono et al., 1977; Wright et al., 1992; Ghazali et al., 2003; P.-C. Kuo et al., 2010; Baidya et al., 1983).

Following the first isolation of CV-A24v from an AHC outbreak in Singapore in 1970, its circulation remained limited to India and South-East Asia until 1985 (Pons-Salort et al., 2015). From this point, CV-A24v spread worldwide, and has now been responsible for two pandemics and numerous outbreaks (Miyamura et al., 1988; Oh et al., 2003). Affected countries include Cuba (Fonseca et al., 2012), China (B. Wu et al., 2014), India (Shukla et al., 2013), Japan (Miyamura et al., 1988) and Thailand (Chansaenroj et al., 2015). Four genotypes (G-I to G-IV) have been identified for CV-A24v. Isolates from the 1970s contain G-I and G-II, those from the 1980s and early 1990s (the first pandemic) contain G-III. Isolates from the second pandemic and more recent outbreaks contain G-IV (Pons-Salort et al., 2015). Such observations suggest that new genetic variants can spread very quickly and infect large populations but are then gradually superseded by new variants.

In general, AHC outbreaks occur in regions with a tropical climate (e.g. Thailand, Singapore, Brazil), with outbreaks usually occurring during the hot rainy seasons. It has been shown that neutralising antibodies produced by infected individuals only provide protection for a few years (Aoki and Sawada, 1992; Goh et al., 1990). Such observations may explain recent studies in Thailand which demonstrated that CV-A24v has been continuously circulating in the country since 2002. During this time, intermittent cases of AHC are reported and large epidemics occur every two to three years, presumably when herd immunity falls below a certain level (Chansaenroj et al., 2015).
Figure 4-10 – Global outbreaks of CV-A24v. Global map showing the geographic origins of CV-A24 isolates for which complete VP1 sequences are available. Non-AHC causing isolates are shown in red and those responsible for AHC are shown in purple, light blue and blue. The circle sizes are proportional to the number of isolates. Adapted from (Baggen, Hurdiss, et al., 2018).
4.1.5 Therapeutic strategies

The highly infectious nature of CV-A24v enables its rapid spread through communities, which places a large burden on education and healthcare systems and can lead to significant economic losses through missed work days (Finger, 2003). At present there are no direct acting antiviral therapies available to control outbreaks of AHC caused by CV-A24v. Efforts to limit the spread of AHC are restricted to control measures such as increased hygiene (frequent hand washing), and advanced warning of outbreaks to members of the community through various communication channels. A number of antiviral compounds have been identified which target various stages of the enterovirus lifecycle (Baggen, Thibaut, et al., 2018), a summary of these is included below.

4.1.5.1 Capsid binders

Capsid binding compounds work by replacing the hydrophobic pocket factor situated at the base the enterovirus canyon, which stabilises the virion and prevents genome uncoating (De Palma et al., 2008; Thibaut et al., 2012). The capsid binding compounds, pleconaril, pirodavir and vapendavir, have proved unsuccessful in clinical trials against rhinoviruses infection (van der Linden, Wolthers, et al., 2015; L. Bauer et al., 2017). However, pleconaril has been reported to improve survival rates in cases of neonatal sepsis caused by non-polio enteroviruses and has undergone a phase II clinical trial (Abzug et al., 2016). A limitation of capsid-binding compounds is the rapid emergence of resistance mutations. Such mutations have been identified through in vitro experiments but also from patients taking pleconaril in clinical trials for rhinovirus infection (Pevear et al., 2005). Whilst such observations raise concerns on the potential clinical application of these compounds, next generation capsid binders have been developed which have a higher barrier to resistance (De Colibus et al., 2014; Kelly et al., 2015).

4.1.5.2 Protease inhibitors

The 3C protease is more conserved than 2A, and as such, has been the primary target for the development of protease inhibitors. The potent peptidomimetic compound rupintrivir has broad spectrum anti-enterovirus activity and functions by binding to the active site of 3C in an irreversible manner (Matthews et al., 1999; Binford et al., 2005; Patick et al., 2005). Despite its poor oral bio-availability,
rupintrivir was taken forward for clinical trials for rhinovirus infection (Hsyu et al., 2002; K.E. Zhang et al., 2001). Whilst this compound appeared to be efficacious for experimentally-infected individuals (Hayden et al., 2003), no therapeutic benefit was observed for individuals suffering from natural rhinovirus infections (L. Bauer et al., 2017). Although further clinical development of rupintrivir has been halted in light of these results, a number of derivatives of rupintrivir are presently in the development stage (J. Tan et al., 2013; Y.W. Tan et al., 2016; C. Wu et al., 2016). In addition, non-peptidomimetic inhibitors with improved bioavailability have been developed but have not yet been subjected to clinical evaluation (G.-H. Ma et al., 2016).

4.1.5.3 2C inhibitors

The 2C protein is a particularly attractive target for the development of antivirals due to its high level of sequence conservation. The 2C protein is an AAA+ ATPase with a number of proposed functions within the virus lifecycle, including helicase activity, reorganisation of cellular membranes, and encapsidation (van der Linden, Wolthers, et al., 2015). A number of structurally-unrelated molecules have been shown to target 2C, one such example is guanidine hydrochloride (De Palma et al., 2008). Furthermore, several FDA-approved drugs have been identified as 2C inhibitors but their mechanism of action is not understood (Ulferts et al., 2016; Ulferts et al., 2013; Zuo et al., 2012). For example, members of the enterovirus B and D species are sensitive to the antidepressant drug fluoxetine (Prozac) (Baggen, Thibaut, et al., 2018; L. Bauer et al., 2017). In addition, fluoxetine was successfully used in a clinical setting to treat an immunocompromised child with chronic enteroviral encephalitis (Gofshteyn et al., 2016). This demonstrates the potential of targeting 2C as an antiviral therapy. Future development of these compounds into effective anti-enteroviral drugs will likely be facilitated by structural elucidation of the 2C protein. A partial crystal structure of EV-A71 2C has recently been reported (H. Guan et al., 2017).

4.1.5.4 Polymerase inhibitors

A number of nucleotide/nucleoside inhibitors (NIs) which target the enterovirus 3D polymerase have been identified. The inhibitory action of these compounds can result from their incorporation into the viral genome or by indirect mechanisms (Jordheim et al., 2013). A number of NIs developed against other
viruses have been shown to have anti-enteroviral activity, such as ribavirin (Pfeiffer and Kirkegaard, 2003; Crotty et al., 2000) and NITD008 (Deng et al., 2014) which were developed against hepatitis C virus (Manns et al., 2001; Fried et al., 2002) and dengue virus respectively. A NI developed as an anticancer drug, gemcitabine, has also been demonstrated to inhibit enteroviruses (Kang et al., 2015). The anti-enteroviral activity of gemcitabine is proposed to result from its incorporation into viral RNA, blocking entry of nucleotides into the 3D active site and by inhibition of ribonucleotide reductase (L. Bauer et al., 2017; Kang et al., 2015).

In addition to NIs, a number of non-nucleoside/nucleotide inhibitors (NNIs) of 3D have been identified (van der Linden, Wolthers, et al., 2015). One such example is GPC-N114 which has broad anti-enteroviral activity and works by targeting the core of 3D where it interferes with template-priming. To date, GPC-N114 is the only known compound which has this mechanism of action, but its efficacy in animal models is yet to be tested (van der Linden, Vives-Adrián, et al., 2015). Another known NNI which targets enteroviruses is amiloride, which reduces the fidelity of the 3D polymerase (van der Linden, Wolthers, et al., 2015; L. Bauer et al., 2017).

4.1.6 Project aims

It has been shown previously that CV-A24v utilises sialic acid as a cellular receptor (E.C. Nilsson et al., 2008). Subsequently, crystallographic studies revealed the structure of CV-A24v and determined the molecular basis of its interaction with sialic acid (Zocher et al., 2014). Nevertheless, several unanswered questions regarding this pandemic human pathogen remain, namely:

- What is the uncoating receptor for CV-A24v and how does this bind to the capsid?
- What do the CV-A24v uncoating intermediates look like?
- Is the process of CV-A24v uncoating a viable target for inhibition with capsid binding compounds and thus a potential therapeutic strategy for AHC outbreaks?

In more general terms, this project aimed to demonstrate the utility of cryo-EM for studying the binding of protein receptors to viruses at better than 4 Å resolution.
The results presented in this chapter are the product of a large collaborative effort between the Universities of Leeds, Utrecht, Tübingen, Umeå and the Spanish National Centre for Biotechnology. This section focusses primarily on the cryo-EM experiments performed by the candidate but complementary functional and crystallographic data is included. Data provided by collaborators is clearly indicated.

4.2 Results

In this study, the early events in the replication cycle of CV-A24v are structurally characterised. ICAM-1 was identified as an essential proteinaceous receptor for CV-A24v using neutralisation assays and confirmed by CRISPR-Cas9 knockout of the ICAM-1 gene. To determine the interaction of ICAM-1 and CV-A24v, single-particle cryo-EM analysis was performed on a complex between a recombinant D1-D2 fragment of ICAM-1 and purified CV-A24v virions. The resulting 3.9 Å structure allowed identification of residues responsible for ICAM-1 binding and demonstrate that this interaction is driven primarily by electrostatic complementarity (Baggen, Hurdiss, et al., 2018). Furthermore, cryo-EM was used to investigate the structure of infectious CV-A24v, together with the structure of an expanded CV-A24v particle that is formed in response to sub-stoichiometric amounts of the natural host-cell receptor ICAM-1. The expanded particle has large pores around its two-fold axis and has lost both its single-stranded RNA genome and the minor capsid protein VP4. Critically, the structure of the empty particle also appears to have lost its pocket factor, a branched ceramide-type lipid that is clearly visible in the native virion structure. The loss of the pocket factor is accompanied by a collapse of the pocket which allows the conformational rearrangement that creates the two-fold pore. This suggests that stabilisation of the pocket with small molecules might inhibit CV-A24v infection. Finally, it was shown that the known capsid binding drug pleconaril can effectively inhibit CV-A24v infection in cell culture, and using macromolecular X-ray crystallography, binding of pleconaril to CV-A24v was shown. Rearrangements were observed in the capsid which allow the linear drug molecular to fit into a pocket that has evolved to bind a branched pocket factor.
4.2.1 ICAM-1 Is an Essential CV-A24v Receptor

CV-A24v was previously shown to bind sialic acid, but this didn't initiate uncoating (Zocher et al., 2014), suggesting that CV-A24v employs an additional, previously-unknown, receptor. To identify this receptor, neutralisation assays were performed with a CV-A24v clinical isolate in Sialic acid-deficient HAP1 CMAS-KO cells (Baggen et al., 2016), using soluble receptor domains (VLDL-R) or antibodies targeting various known enterovirus receptors (ICAM-1, DAF, integrin α2, PVR, CAR, integrin αvβ3, and PSGL-1) (Figure 4-11A). The ability of these ligands to suppress receptor binding was confirmed using positive control viruses (Baggen, Hurdiss, et al., 2018). Only the antibody against ICAM-1, a cell adhesion molecule consisting of five Ig-like domains (van de Stolpe and van der Saag, 1996) (Figure 4-11B), blocked CV-A24v infection (Figure 4-11A). To determine whether ICAM-1 is essential for CV-A24v, the ICAM-1 gene was knocked out in various cell types by CRISPR/Cas9 genome editing (Figure 4-11B). The ICAM-1 knockout (ICAM-1 KO) genotype was verified by sequence analysis and by assessing the susceptibility of ICAM-1 KO cells to infection with ICAM-1–dependent coxsackievirus A21 (CV-A21) (Shafren et al., 1997) and rhinovirus B14 (RV-B14) (Greve et al., 1989) and CAR-dependent coxsackievirus B3 (CV-B3) (Bergelson et al., 1997) (Figure 4-13A-B). CV-A24v infection was reduced by ICAM-1 knockout in HAP1 CMAS KO cells, wild-type HAP1 cells, and HeLa cells and was rescued by transfecting ICAM1 cDNA (Figure 4-11C).
Figure 4-11 – Identification of ICAM-1 as a CV-A24v receptor. (A) Yields of infectious virus after a single replication cycle. Sia-deficient HAP1 CMASKO cells were treated with antibodies against enterovirus receptors (ICAM-1, DAF, integrin α2, PVR, CAR, integrin αvβ3, and PSGL-1) or soluble receptor (VLDL-R) and infected at 37 °C with CV-A24v 110387. (B) Schematic representation of ICAM-1 showing Ig-like domains 1–5 (D1–D5) and the two sites targeted by CRISPR/Cas9 (gRNA 1 and 2). Knockout was accomplished by disrupting a 1,489-bp region of ICAM1 encoding the transmembrane domain (TM), either by excision of this region or by introducing a frameshift mutation. (C) HeLa-R19 ICAM-1KO transfected with plasmid encoding ICAM-1 cDNA were exposed to virus and yields of infectious virus were measured after a single replication cycle. Adapted from (Baggen, Hurdiss, et al., 2018). Experiments performed by Jim Baggen and Dr Hendrik Thibaut, Utrecht University.
To determine if ICAM-1 directly interacts with CV-A24v biolayer interferometry was used. This demonstrated that CV-A24v binds directly to this receptor when compared to ICAM-1 utilising RV-B14 and ICAM-1 independent CV-B3 (Figure 4-12A). The necessity for a direct interaction between CV-A24v and ICAM-1 was further demonstrated using a neutralisation experiments where the presence of soluble ICAM-1 ectodomain was shown to inhibit infection of primary conjunctival cells (Figure 4-12B)

**Figure 4-12 – ICAM-1 interacts directly with CV-A24v.** (A) Biolayer interferometry analysis of virus binding to either ICAM-1 or negative control receptor LAMP-1. (B) Relative levels of CV-A24v 110390 bound to primary conjunctival cells with expanded lifespan (HC0597), in the presence of ICAM-1 D1D5 or negative control receptor CAR D1. Adapted from (Baggen, Hurdiss, et al., 2018). Experiments performed by Dr Eric de Vries and Dr Xander de Haan, Utrecht University.

Analysis of virus production in HeLa cells (Figure 4-13A) and analysis of cytopathic effect induction in HeLa (Figure 4-13B) or human corneal epithelial (HCE) cells showed that several CV-A24v clinical isolates, the nonvariant prototype strain Joseph, and a nonvariant clinical isolate all require ICAM-1 for infection. Sialylation of ICAM-1 was not found to be essential for interaction with CV-A24v and CV-A24 Joseph, as shown by biorlayer interferometry (data not shown). Taken together, these data indicate that ICAM-1 functions as an essential receptor for AHC-causing variant strains as well as nonvariant strains.
**Figure 4-13 – ICAM-1 is essential for variant and non-variant CV-A24.** (A) HeLa-R19 cells were infected and yields of infectious virus were determined after a single replication cycle. Dashed lines represent virus input levels (T = 0). Error bars represent the mean ± SEM of three to four biological replicates. (B) HeLa-R19 cells were infected with seven CV-A24v clinical isolates, two CV-A24 nonvariant strains and ICAM-1-dependent (CV-A21, RV-14) or CAR-dependent (CV-B3) control viruses, followed by crystal violet staining of surviving cells. Adapted from (Baggen, Hurdiss, et al., 2018). Experiments performed by Jim Baggen and Dr Hendrik Thibaut, Utrecht University.

### 4.2.2 The Structure of CV-A24v in Complex with ICAM-1

To elucidate the molecular basis of the interaction between CV-A24v and ICAM-1, it was decided to use cryo-EM to analyse a complex between purified virions and a recombinant D1-D2 fragment of ICAM-1. An initial attempt to mix the virus and receptor in solution resulted in rapid precipitation of the sample. Precipitated material was removed by centrifugation and cryo-EM grids were prepared with the supernatant. No virions could be identified when these grids were examined by cryo-EM. To mitigate this effect, an on-grid binding protocol was employed (Drulyte et al., 2018). Briefly, CV-A24v virions were immobilised on a lacey carbon support and left for 30 seconds before excess material was blotted away. Immediately after this, D1-D2 was added for 30 seconds at 8 °C before blotting and plunge freezing (Figure 4-14A).
**Figure 4-14 – Data collection and image processing of CV-A24v: ICAM-1.** (A) Typical micrograph of CV-A24v in complex with ICAM1-D1D2 (Scale bar, 50 nm). (B) Representative 2D class averages generated in Relion. (C) Central slice through 3D reconstruction generated in Relion. (D) A plot of the Fourier shell coefficient (FSC). Based on the 0.143 criterion for the gold standard comparison of two independent datasets, the resolution of the reconstruction is 3.9 Å. Adapted from (Baggen, Hurdiss, et al., 2018).

This rapid on-grid binding approach allowed the cryo-EM structure of CV-A24v bound ICAM-1 D1-D2 to be determined (Figure 4-14B-C). It was speculated that the longer incubation in solution resulted in loss of particle integrity, presumably due to receptor-mediated destabilisation of the particle. The global resolution of the EM density is 3.9 Å (Figure 4-14D), the resolution of the viral capsid proteins is ~3.6–3.8 Å (Figure 4-15A-B), and that of the capsid-interfacing region of ICAM-1 D1 is ~3.7 Å (Figure 4-15D), sufficient to resolve bulky side chains in the electron density map (Figure 4-16C). The D2 domain of ICAM-1 is poorly resolved in the map with a local resolution estimate of worse than 10 Å (Figure 4-15C).
This indicates that this region of the molecule is highly flexible as described previously (Xiao et al., 2005).

Figure 4-15 – Local resolution of the CV-A24v-ICAM-1 map. (A) The cryo-EM reconstruction of CV-A24v in complex with ICAM-1 (D1D2) viewed down the icosahedral two-fold axis and coloured according to local resolution (3.6σ). (B) A 40-A slab through the centre of the map shown in A. (C) The cryo-EM reconstruction of CV-A24v in complex with ICAM-1 (D1D2) viewed down the icosahedral two-fold axis and coloured according to local resolution (0.6σ). (D) Segmented D1 domain from ICAM-1 coloured according to local resolution. Adapted from (Baggen, Hurdiss, et al., 2018).
The outer surface of an enterovirus particle is composed of sixty copies of the structural proteins VP1, VP2, and VP3 (Figure 4-16B). A depression encircling each five-fold axis of symmetry forms the so-called canyon, which is the binding site for all uncoating receptors identified to date. The structure shows that the D1 domain binds in the canyon at the quasi-three-fold axis of the capsid (Figure 4-16A). In the final reconstruction, the signal relating to the capsid interfacing region of ICAM-1 D1 is approximately 60% as strong as that of the adjacent capsid region, indicating 60% occupancy of the receptor binding site.

**Figure 4-16** – The structure of CV-A24v in complex with ICAM-1. (A) Enlarged tilted view of density denoted by a black triangle in B, containing the corresponding atomic model. ICAM-1 D1 binds in the “canyon” located at the quasi-three-fold axis. (B) The cryo-EM reconstruction of CV-A24v in complex with ICAM-1 (D1D2) viewed down the icosahedral two-fold axis. VP1-3 (3.6σ), D1 (1σ), and D2 (0.6σ). (C) Typical example of the cryo-EM electron density of VP1-3 and D1 domain. Adapted from (Baggen, Hurdiss, et al., 2018).
ICAM-1 engages the canyon of CV-A24v in a manner similar to that of CV-A21. In both cases, the buried surface area of this interaction is approximately 30% less than that observed for ICAM-1 in complex with HRV-14 and HRV-16. In contrast to earlier studies of ICAM-1 binding to CV-A21, HRV-14 and HRV-16 (Xiao et al., 2001; Xiao et al., 2005; Kolatkar et al., 1999; Olson et al., 1993), the interface between the CV-A24v capsid and ICAM-1 is well defined in the final EM density map, enabling secondary structure and the majority of side chains to be resolved in this portion of the structure. This, in turn, allowed calculation of a detailed footprint of the D1 component of the receptor on the virus surface, shown as a “roadmap” projection (Figure 4-17).
Figure 4-17 – Receptor footprint of ICAM-1. Radially coloured isosurface representation of CV-A24v in complex with D1 (grey) viewed down the icosahedral two-fold axis (Left). The stereographic projection of the viral surface (Right), where the polar angles $\theta$ and $\varphi$ represent latitude and longitude, respectively. Amino acids which interact with ICAM-1 D1 are located at the floor and wall of the canyon (blue) and are circled in white. The radial colouring key is shown in ångströms. Adapted from (Baggen, Hurdiss, et al., 2018).
The specificity in the interaction between ICAM-1 and CV-A24v is largely driven by electrostatic complementarity (Figure 4-18A and Figure 4-19). Specifically, Lys\textsuperscript{29} (FG loop) of ICAM-1 interacts with VP3 Asp\textsuperscript{181} and VP1 Asp\textsuperscript{238} (Figure 4-18B). In addition, Asp\textsuperscript{71} on the tip of the FG loop of ICAM-1 forms a salt bridge with VP1 Arg\textsuperscript{168}. Finally, Lys\textsuperscript{50} (strand C) and Lys\textsuperscript{39} (strand D) of ICAM-1 form salt bridges with VP1 Asp\textsuperscript{228} and Asp\textsuperscript{225} (GH loop).

Figure 4-18 – Salt bridges involved in the CV-A24v–ICAM-1 interaction. (A) Surface representation of the EM-derived atomic model for the quasi-three-fold axis of CV-A24v (grey) and ICAM-1 D1 (pink), with residues forming salt bridges labelled and coloured in respect to their interacting partners. (Left) The surface of CV-A24v. (Right) The surface of ICAM-1 D1 that interacts with CV-A24v. Residues forming salt bridges are coloured according to charge (red is negative, blue is positive) and capsid residues that hydrogen-bond with ICAM-1 are dark grey. (B) Overview of electrostatic interactions shown in E between ICAM-1 (pink) and CV-A24v (blue) with potential hydrogen bonds denoted as dotted black lines and EM density as a green mesh. Adapted from (Baggen, Hurdiss, et al., 2018).
Figure 4-19 – Surface electrostatics of CV-A24v and ICAM-1. Electrostatic surface representation of the CV-A24v capsid (Left) and interacting face of ICAM-1 D1 (Right) coloured red to blue ranging from most negative (−10 kT/e) to most positive (+10 kT/e). Neutral surface is coloured white. The interacting surfaces are indicated by a dashed line. Adapted from (Baggen, Hurdiss, et al., 2018).

Additional interactions include a hydrogen bonding network that tethers the DE loop of ICAM-1 D1 (Arg⁴⁹, Asn⁴⁷ and Pro⁴⁵) to VP1 (Tyr²³⁰) and VP2 (Ala¹³⁸, Lys¹³⁹ and Thr¹⁴⁰) (Figure 4-20B). Many of these residues are conserved among other ICAM-1–binding viruses belonging to the enterovirus C species, suggesting a conserved mode of interaction with ICAM-1 (Xiao et al., 2005; Newcombe et al., 2003). On the contrary, only VP1 Asp²³⁸ and VP3 Asp¹⁸¹ are present in ICAM-1–binding rhinoviruses. In addition, there appears to have been a lack in amino acid diversity in the ICAM-1 binding site of CV-A24v compared to nonvariant strains. The model was then compared to that of the crystal structure of CV-A24v (PDB 4Q4Z) - (Zocher et al., 2014) and the N-terminal two domains of ICAM-1 (1IAM) - (Bella et al., 1999). This revealed that binding of ICAM-1 one to CV-A24v induces a conformational change in the VP2 puff region and the GH loop of VP1. Furthermore, the DE loop of ICAM-1 in the cryo-EM derived model exhibits structural differences to that of the reported crystal structure (Figure 4-21).
Figure 4-20 – Hydrogen bonding interactions between ICAM-1 and CV-A24v.
(A) Ribbon diagram of ICAM-1 D1 monomer in chainbow colouring with secondary structure labelled as described previously (9). (B) Hydrogen bonding network between the DE loop of D1 (grey) - corresponding to the boxed region in (A), VP2 (green), and VP1 (blue) with hydrogen bonds shown as dashed black lines. Oxygen and nitrogen atoms are coloured red and blue, respectively. Adapted from (Baggen, Hurdiss, et al., 2018).
### Table 4-1 - Amino acid composition of CV-A24 capsid sites interacting with ICAM-1. Adapted from (Baggen, Hurdiss, et al., 2018).

<table>
<thead>
<tr>
<th>Type of interaction</th>
<th>Protein</th>
<th>Location</th>
<th>Nonvariant</th>
<th>First AHC-causing variant (EH24/70)</th>
<th>Variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salt bridges</td>
<td>VP1</td>
<td>168</td>
<td>Q/R/K/I</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td>225</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>228</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>238</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>VP3</td>
<td>181</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Hydrogen bonds</td>
<td>VP1</td>
<td>223</td>
<td>T/A/D/V</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>224</td>
<td>T/A</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td>230</td>
<td>F/Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td></td>
<td>284</td>
<td>E/D</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>VP2</td>
<td>138</td>
<td>A/N/V</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>139</td>
<td>K</td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td></td>
<td></td>
<td>140</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>142</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>VP3</td>
<td>182</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
</tbody>
</table>
Figure 4-21 – Conformational changes in CV-A24v and ICAM-1. Ribbon diagram of ICAM-1 D1 domain (grey) bound to CV-A24v VP1 (blue) and VP2 (green) derived from our 3.9 Å resolution cryo-EM map. The equivalent view from the crystal structure of CV-A24v (4Q4W) and ICAM-1 (1IAM) are overlaid and shown transparent.

4.2.3 Uncoating of CV-A24v

In the CV-A24v-ICAM-1 data set, a small number of empty virions were evident in the micrographs (Figure 4-14A). Single-particle image processing of this subset of empty virions (951 particles) yielded a 21 Å resolution reconstruction (data not shown) with insufficient structural details to warrant further interpretation. To further investigate the abundance of empty virions within the purified sample, a second data collection was performed on what was believed to be a CV-A24v only cryo-EM grid prepared at the same time as those described previously (Baggen, Hurdiss, et al., 2018). To improve the signal to noise ratio, this second data set was collected on an energy-filtered, K2 summit direct electron detector operated in super-resolution mode. As before, empty virions were apparent in the micrographs (Figure 4-22A). In addition, a number of these empty virions were observed directly adjacent to diffuse, electron dense material, consistent with the viral genome in the process of, or having just egressed from
the capsid (Figure 4-22B). The full and empty particles were individually processed in Relion 2.0 which allowed their structures to be determined at 3 Å and 3.8 Å respectively (Figure 4-23A-B). The resolution achieved for the empty particle reconstruction is particularly impressive given that it contained only 1206 particles.

![Figure 4-22 – Empty CV-A24v virions in micrographs. (A) Cryo-EM micrograph of CV-A24v virions in vitreous ice showing both full and empty particles. (B) Enlarged view of example empty virions with electron dense material immediately adjacent to one side of the capsid (top). On the bottom panel, the capsid is coloured blue and the additional material is coloured red.](image-url)
Completely unexpectedly, weak density for ICAM-1 was observable in the full CV-A24v structure, despite supposedly not being present in the grid. In order to rule out contamination of the purified CV-A24v stocks, silver staining and mass spectrometry analysis was performed (data not shown) which didn’t detect any trace of ICAM-1. To further control for this observation, new CV-A24v grids were prepared, without ICAM-1, and a large data set of 11,000 micrographs was collected. From this data, a 3.2 Å resolution reconstruction of the full particle was generated. Importantly, no ICAM-1 density was present in this reconstruction which led to the conclusion that ICAM-1 had been introduced previously through user error during cryo-grid preparation. It was apparent from a comparison of the three full CV-A24v reconstructions that ICAM-1 was present in a lower quantity when compared to the previous 3.9 Å resolution reconstruction (Figure 4-24).
Therefore, the structures are referred to as fully decorated (FD), sub-stoichiometrically decorated (SSD) and undecorated (UD), respectively.

![Comparison of UD, SSD and FD reconstructions](image)

**Figure 4-24 – Comparison of UD, SSD and FD reconstructions.** Central slices through the undecorated, sub-stoichiometrically decorated and fully decorated CV-A24v reconstructions.

As with the FD and SSD data sets, a number of empty particles were observed in the UD micrographs. From these it was possible to generate a 4.3 Å reconstructions of the empty particle and A-particle (Figure 4-25). The empty particle and A-particle contained 1218 and 1046 particles respectively. Whilst a similar overall number of empty particles were observed in the UD, SSD and FD data sets (~1000), they were proportionally more abundant in the ICAM-1 containing micrographs, ~3-4% vs ~1%. In addition, whilst A-particles also accounted for ~1% of the UD particles, they were not found in either the FD or SSD data. From these observations, it was speculated that ICAM-1 may be able to trigger uncoating of CV-A24v, even on the 30 second incubation period required for the on-grid binding protocol used to avoid aggregation in solution. Therefore, the inadvertent introduction of ICAM-1 into the SSD dataset may have
been serendipitous in producing enough empty particles to generate the 3.8 Å reconstruction.

<table>
<thead>
<tr>
<th>Undecorated (UD)</th>
<th>(Altered) A-particle</th>
<th>Empty particle</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2 Å</td>
<td>4.3 Å</td>
<td>4.3 Å</td>
</tr>
</tbody>
</table>

![Images of particles](image)

**Figure 4-25 – Full, altered and empty CV-A24v particles in the UD data.**
Central slices though the full, altered particle and empty particle reconstructions generated from the undecorated CV-A24v data.

Local resolution estimation indicated that the majority of the SSD CV-A24v structure was at 2.8 Å resolution (Figure 4-26A), and several of the ordered waters from the CV-A24v crystal structure (PDB 4Q4W) were visible (Zocher et al., 2014). Furthermore, strong density was observed in the hydrophobic pocket for a branched pocket factor which could in principle accommodate part of a ceramide molecule (Figure 4-28). The residues making hydrophobic interactions with the pocket factor are shown in Figure 4-38. The local resolution of the empty map was between 3.6-4.0 Å resolution (Figure 4-27), sufficient to identify bulky sidechains within the map and permit atomic modelling (Figure 4-27).
**Figure 4-26 – Local resolution of the full and empty CV-A24v maps.** Local resolution of the full (A) and empty (B) CV-A24v structures. A single asymmetric unit (seen from outside the capsid) is shown coloured according to its local resolution in both cases. The interior of the asymmetric unit is also shown by rotating the asymmetric unit 180° about the y-axis. The highest resolution bin is 2.8 Å (blue) and the lowest resolution bin is 4.4 Å (red). A key is shown for reference.
Figure 4-27 – Map quality of the full and empty CV-A24v reconstructions. Typical example of the cryo-EM electron density of VP1-4 for the SSD full CV-A24v structure and the equivalent views of VP1-3 for the empty structure.
As seen for expanded particles of other enteroviruses, the empty CV-A24v capsid is ~4% larger than the full particle and has adopted a more facetted, angular appearance (Figure 4-29A-B). Structural features characteristic of expanded enterovirus particles are apparent, including an opening across the two-fold symmetry axis, with dimensions of 36.6 Å by 8.4 Å, and at the quasi-three-fold, with dimensions of 24.6 Å by 11.2 Å. These global structural changes arise from tectonic movements in the particle which resulting in separation of the asymmetric units. Furthermore, the internal coat protein, VP4, is lost (Figure 4-29C-D) and weak EM density is observed protruding from the quasi-three-fold opening (Figure 4-29E). The resolution of this density did not permit atomic modelling but it was hypothesised that this corresponds to the exposed N-terminus of VP1, described previously for CV-A16 (Ren et al., 2013). As expected, the structure of the CV-A24v empty particle is very similar to other expanded enterovirus particles (Xu et al., 2017; Ren et al., 2013; X. Wang et al., 2012). Comparison of the Cα positions between the CV-A24v empty particle and A-particles described for CV-A6, CV-A16 and EV-A71, gave RMSD values of

**Figure 4-28 – The CV-A24v pocket factor.** (A) A single asymmetric unit from the SSD full CV-A24v structure with the pocket factor coloured orange. (B) Enlarged view of the hydrophobic pocket showing density for VP1 and the pocket factor containing the fitted atomic model.
0.96, 0.85 and 0.96, respectively. The opening at the two-fold symmetry axis has a similar dimension in each of these viruses, however, differences are apparent at the quasi-three-fold. In CV-A16 and EV-71, the quasi-three-fold opening is more obstructed compared to CV-A6. The analogous region in CV-A24v more closely resembles that of CV-A6 and is arguably even less obstructed (Figure 4-30). Other prominent structural difference between these enteroviruses localise to the four surface loops of VP1 (BC, DE, EF, and HI) – data not shown.
Figure 4-29 – Global conformation changes in the CV-A24v capsid. (A) Cryo-EM reconstruction of the full CV-A24v particle viewed down the icosahedral two-fold axis (ICAM-1 density not shown). (B) Equivalent view of that shown in (A) for the empty particle, demonstrating the particle expansion and opening at the icosahedral two-fold and quasi three-fold. (C) Internal view of the full particle showing the VP4 protein (yellow). (D) Equivalent view as (C) for the empty particle. (E) Zoomed-in view of the icosahedral two-fold of the empty CV-A24v. Additional unmodeled density proposed to be the VP1 N-terminus, is shown as a blue mesh emerging from the quasi three-fold opening.
Figure 4-30 – Comparison of CV-A24v empty particle to expanded capsids of other enteroviruses. Close up views of the two-fold symmetry axis (black ellipses) of (A) CV-A24v empty particle. (B) CV-A6 A-particle (PDB: 5XS4) (Xu et al., 2017). (C) CV-A16 (PDB: 4JGY) (Ren et al., 2013). (D) EV-A71 (PDB: 4N43) (Lyu et al., 2014). The quasi-three-fold channel is indicated by dashed ellipses.
A number of local structural differences are apparent between the full and empty CV-A24v particle, with a calculated RMSD value of 1.37 between the two aligned asymmetric units. In the empty particle, the hydrophobic cavity is collapsed to a state which is incompatible with pocket-factor binding, from a volume of 217 Å³ in the full virus, to 26 Å³ (Figure 4-31). This collapse in the pocket arises through small conformational changes it walls. The largest movement of the polypeptide backbone is in VP1, in residues 237-243 which are at the end of the GH loop and start of strand H, where Cα movements approach ~3 Å. This causes Phe²³⁹ to change its position and ablate the pocket. In addition, VP1 strand E also exhibits Cα movements of ~2 Å which causes Tyr¹⁶¹ to move inwards into the pocket and its sidechain swings ~90° to occlude the hydrophobic cavity (Figure 4-32A). In the full CV-A24v particle, Phe²³⁹ and Tyr¹⁶¹ both participate in hydrophobic interactions with the branched pocket factor. Specifically, Phe²³⁹ is wedged between the two branches of this lipid moiety, apparently acting as a molecular spacer (Figure 4-32). Therefore, the movements seen in these two side chains in the empty particle seem to be the major determinant of hydrophobic pocket collapse (Figure 4-32C). A number of regions which are well resolved in the full particle are disordered in the empty counterpart. Residues 212-236 in the VP1 GH loop and 133-146 in the VP2 puff region are unresolved in the empty particle (Figure 4-33). Given that these undergo conformational change upon ICAM-1 binding (Figure 4-21), it seems likely that this disorder arises from strain induced upon binding of the protein receptor. The GH loop is directly upstream of strand H and has previously been termed the sensor-adaptor, further suggesting that ICAM-1 is the uncoating receptor for CV-A24v. The HI loop located at the five-fold symmetry axis of the particle forms part of the sialic acid binding site for CV-A24v (Zocher et al., 2014). This loop is also disordered in the empty particle, suggesting that empty and A-particles should not be able to engage either receptor.
Figure 4-31 – Collapse of the CV-A24v hydrophobic pocket. The atomic coordinates for two adjacent asymmetric units from the (A) full and (B) empty CV-A24v structure (grey). The solvent-accessible volume (orange) of the hydrophobic pockets from both models was computed using CASTp 3.0 (Tian et al., 2018).
**Figure 4-32 – Mechanics of the hydrophobic pocket collapse.** (A) Aligned atomic coordinates of the full and empty CV-A24v structures, showing a close up view of the hydrophobic pocket region. (B) EM density and fitted model for the hydrophobic pocket in the full CV-A24v particle with residues coloured blue and the pocket factor coloured orange. (C) Equivalent view to that shown in (B) for the empty particle.
Figure 4-33 – Disorder regions in the CV-A24v empty particle. (A) Atomic coordinates of an asymmetric unit from the full CV-A24v structure. (B) Equivalent view as shown in (A) for the empty particle. The coordinates for bound sialic acid (PDB: 4Q4Z) (Zocher et al., 2014) and ICAM-1 are overlaid in both structures for comparison. Dashed lines in (B) indicate regions which become disordered upon particle expansion.
4.2.4 Inhibition of CV-A24v

Although the conformational changes associated with uncoating of CV-A24v had now been characterised, it remained uncertain whether this process could be inhibited with known capsid binding compounds. The hydrophobic pocket of CV-A24v was first compared to those of other enteroviruses which are a known to be targeted by capsid binding compounds. Interestingly, Tyr$^{161}$ in the pocket of CV-A24v adopts a different conformation to the equivalent aromatic residue observed in other enteroviruses (Figure 4-34A). Whilst both possible conformers were modelled in the crystal structure of CV-A24v (Zocher et al., 2014), the 3.0 Å resolution full structure indicated that the atypical conformation of this residue is consistent with the experimental density (Figure 4-34C-D). The presence of a branched pocket factor in CV-A24v makes it unique amongst other human enteroviruses which have been structurally characterised. The only other known enterovirus with a branched pocket factor is the animal pathogen, swine vesicular disease virus (SVDV) (Verdaguer et al., 2003). However, the pocket factor observed in SVDV is rotated ~170° relative to that of CV-A24v and the equivalent residue of Tyr$^{161}$ adopts the conventional conformation observed in other human enteroviruses (Figure 4-34B).
Figure 4-34 – Comparison of enterovirus pocket factors. (A) Comparison of the pocket factor modelled in CV-A24v compared to those of other structurally characterised human enteroviruses. (B) Comparison of the CV-A24v pocket factor to that of SVDV. Enterovirus names are shown and correspond to the following accession codes: CV-A21 (PDB: 1Z7S) (Xiao et al., 2005), EV-D68 (PDB: 4WM8) (Liu et al., 2015), HRV-14 (PDB: 1K5M) (Ding et al., 2002), CV-A16 (5C4W) (Ren et al., 2015), EV-A71 (PDB: 3VBS) (Wang et al., 2012), PV-1 (PDB: 1HXS) (Miller et al., 2001), CV-B3 (PDB: 1COV) (Muckelbauer et al., 1995), SVDV (1MQT) (Verdaguer et al., 2003). (A) EM density from the full CV-A24v reconstruction containing the atomic coordinates from the previously determined crystal structure (PDB: 4Q4W). (B) As shown in (A), following removal of the Tyr^{161} conformation which is not supported by the EM density.
The hydrophobic pocket of CV-A24v was next compared with that of other human enteroviruses bound by capsid binding compounds. In each case, the atypical conformation of Tyr\(^{161}\) appears as if it would preclude binding of these compounds to CV-A24v (Figure 4-35). On the basis of this observation, it was hypothesised that the structurally characterised strain of CV-A24v from this study may be resistant to capsid binding compounds. To test this, we incubated purified CV-A24v virions with an increasing concentration of the original capsid binder, pleconaril, and determined its effect by endpoint serial dilution (Figure 4-36) and immunofluorescent staining (data not shown). Surprisingly, despite having a seemingly incompatible hydrophobic pocket for capsid binders, pleconaril was able to robustly inhibit CV-A24v, with an EC\(_{50}\) value of ~500 nM.

\[\text{Figure 4-35 – Comparison of inhibitor bound enteroviruses.} (A) \text{Comparison of hydrophobic pockets of enterovirus bound by pleconaril compared to the full CV-A24v model.} (B) \text{Comparison of hydrophobic pockets of enterovirus bound by GPP3 compared to the full CV-A24v model.} \text{Enterovirus names are shown and correspond to the following PDB accession codes: EV-D68 (PDB: 4WM8) (Liu et al., 2015), CV-A21 (PDB: 1Z7S) (Xiao et al., 2005), HRV-14 (PDB: 1NCQ) (Zhang et al., 2004), CV-A16 (PDB: 5ABJ) (De Colibus et al., 2015), EV-A71 (PDB: 4CDU) (De Colibus et al., 2014), PV-3 (PDB: 5O5P) (Marsian et al., 2017)}\]
To investigate how pleconaril binds to a seemingly incompatible hydrophobic pocket, X-ray crystallography was used to determine the structure of CV-A24v in complex with the pleconaril at 1.8 Å resolution. This revealed strong density for the bound drug molecule in the hydrophobic pocket. Binding of the pleconaril does not induce major changes in the structure of CV-A24v, with an RMSD value of 0.26 between the pocket factor and pleconaril-bound VP1 molecules (Figure 4-37A). The Cα movements reach ~0.6 Å in the residues 227-230 of the GH loop of VP1 which interact with ICAM-1 but it is unclear whether such small conformational change would preclude binding to the receptor molecule. Pleconaril is able to bind deeper into the hydrophobic pocket of CV-A24v compared to the natural pocket factor (Figure 4-37A). In order to accommodate the capsid binding compound, a number of conformational changes occur in residues lining the hydrophobic pocket. In particular, the sidechain of Tyr^{161} swings ~90°, which allows the drug molecule to reach further into the hydrophobic
core of the VP1. More subtle differences are also observed in additional residues, including Leu\textsuperscript{134}, Met\textsuperscript{136}, Tyr\textsuperscript{207} and Tyr\textsuperscript{239} (Figure 4-37B). There are eight VP1 residues which participate in hydrophobic interactions with both the pocket factor and pleconaril (Figure 4-38). An additional ten residues interact with the second branch of the natural lipid molecule including three VP3 sidechains, His\textsuperscript{108}, Arg\textsuperscript{177} and Leu\textsuperscript{223}, from an adjacent asymmetric unit (Figure 4-38A). In contrast, only two additional residues interact with the pleconaril molecule, including VP3 Ala\textsuperscript{23} from within the same asymmetric unit (Figure 4-38B).
Figure 4-37 – Pleconaril binding to CV-A24v. (A) Aligned VP1 proteins and ligands from the pocket factor and pleconaril-bound CV-A24v structures coloured blue and grey, respectively. An ICAM-1 molecule is shown to indicate the position of receptor binding to the GH loop of VP1, indicated by an asterisk. (B) Overlaid atomic coordinates for the bound pocket factor and pleconaril showing select residues surrounding each molecule. X-ray crystal structure determined by Dr Georg Zocher, University of Tübingen.
Figure 4-38 – Pocket factor and pleconaril binding residues. LigPlot of the residues which make hydrophobic interactions with (A) the CV-A24v pocket factor and (B) pleconaril. VP1 and VP3 residues names are coloured blue and red, respectively. Residues which bind to both the pocket factor and pleconaril are circled (Laskowski and Swindells, 2011).
4.3 Discussion

The structures presented here provide snapshots of the early events in the life cycle of CV-A24v, from receptor binding to genome uncoating. Enteroviruses are known to engage a broad spectrum of receptors which can be classified as attachment or uncoating factors (Rossmann et al., 2002; Baggen, Thibaut, et al., 2018). It was previously reported that CV-A24v engages an α2,6-linked sialic acid moiety at a binding site situated around the five-fold symmetry axis of the capsid. Binding of sialic acid to CV-A24v was not found to trigger uncoating of the viral genome and thus likely functions primarily to mediate attachment to host cells (Zocher et al., 2014). Indeed, it is widely accepted that receptor molecules which bind in regions distal to the canyon serve primarily as attachment factors. Whilst EV-D68 was shown to engage sialic acid in the canyon, triggering ejection of the pocket factor and conformational changes (Yue Liu, Sheng, Baggen, et al., 2015), this is clearly not the case for CV-A24v. It was therefore hypothesised that CV-A24v may engage an additional receptor which mediates uncoating of the viral genome. Using a neutralisation screen and CRISPR-Cas9 knockout, ICAM-1 was identified as an essential receptor for both CV-A24v and non-variant strains. The molecular basis of this interaction was determined using high-resolution cryo-EM. As expected, the receptor binds in the canyon as seen for other ICAM-1 binding enteroviruses such as rhinoviruses (Greve et al., 1989; Xing et al., 2003) and CV-A21 (Xiao et al., 2001; Xiao et al., 2005). Whilst previously reported complexes had resolution ranging between ~7-25 Å, the one presented in this study provides the first high-resolution glimpse of an enterovirus ICAM-1 interaction. Furthermore, this represent the first sub-4 Å resolution structure of an enterovirus in complex with a proteinaceous receptor. The well-resolved capsid and D1 domain of ICAM-1 permitted identification of side-chain positions, thus allowing a detailed roadmap of this interaction to be determined. Excluding potential solute molecules which are not visible at this resolution, the interaction of ICAM-1 and CV-A24v was shown to have a $\Delta G$ value of -1.8 kcal/mol. For CV-A24v, it has been reported that the emergence of the AHC-causing CV-A24v variant was accompanied by a loss of amino acid diversity, probably as a result of a population bottleneck (Smura et al.,
2014). Consequently, many changes in nucleotide sequence exist between variant and nonvariant strains, and a reduction in amino acid diversity occurred at several sites within the CV-A24 capsid that interact with ICAM-1 (Table S3). We speculate that some of these may have caused subtle changes in the engagement of ICAM-1 that facilitated ocular tropism. Thus, the identification of ICAM-1-binding residues in CV-A24v may help future studies investigating whether changes in ICAM-1 binding have contributed to the earliest manifestation of AHC. Furthermore, many of the residues shown to interact with ICAM-1 are conserved amongst other members of the enterovirus C species which use this receptor (Newcombe et al., 2003) and may therefore represent a conserved mode of binding. In contrast, only VP1 Asp$^{238}$ and VP3$^{181}$ are present in rhinoviruses that bind ICAM-1.

Conformational differences are evident in the DE loop of ICAM-1 bound to CV-A24v when compared to the published crystal structure of this molecule alone (Bella et al., 1998). It is tempting to speculate that the conformational changes induced in ICAM-1 upon binding of CV-A24v, trigger endocytosis of the virus-receptor complex. In addition to the receptor molecule, the VP1 GH loop and VP2 puff region, which directly engage ICAM-1, have conformational changes compared to the unbound virion (Zocher et al., 2014). Furthermore, approximately four-fold more empty virions were present in cryo-EM data sets containing ICAM-1. Together these observations suggest that ICAM-1 is the uncoating receptor for CV-A24v. The relatively small increase in the abundance of empty virions likely reflect the short timescale (30 seconds) used for on-grid binding of the receptor, necessary to avoid precipitation of the sample. Going forward, longer incubation periods may show a time-dependent increase in the number of empty CV-A24v particles. For example, such studies on receptor-mediated uncoating of RV3 (Xing et al., 2003) and EV-71 (Dang et al., 2014) were performed for 15 minutes and 4 hours, respectively. The A-particle and empty particle of CV-A24v presented in this section are structurally indistinguishable other than the presence and absence of genome density. As expected the CV-A24v empty particle is morphologically very similar to those of expanded particles of other enteroviruses. The opening at the two-fold symmetry axis of CV-A24v most closely resembles that of CV-A6 (Xu et al., 2017) with both having relatively unobstructed quasi-three-fold
channels. In contrast, this opening is narrower in EV-A71 (X. Wang et al., 2012) and CV-A16 (Ren et al., 2013). It is difficult to ascertain whether the differences observed between these enteroviruses are genuine or result from the different techniques used to determine their structures. One could argue that restrictions imposed though formation of a crystal lattice may permit more flexible regions to be modelled for CV-A16 and EV-A71. Indeed, many regions of the CV-A24v structure become disordered upon particle expansion. The VP1 GH loop, previously termed the sensor adapter (X. Wang et al., 2012), becomes disordered and undergoes conformational changes which result in the release of the pocket factor and collapse of the hydrophobic pocket. Whilst there are few surprises in the global structure of the expanded CV-A24v particle, aspects of the hydrophobic pocket appear to be unique amongst human enteroviruses which have been characterised thus far. In the full virion, density consistent with a branched pocket factor is apparent. Whilst other human enteroviruses have linear lipid molecules, CV-A24v is able to accommodate a branched molecule by adopting an alternative conformation of Tyr\textsuperscript{161}.

It has previously been suggested that enteroviruses with a short pocket factor, such as HRV-16 (Pevear et al., 2005) and EV-D68 (Yue Liu, Sheng, Fokine, et al., 2015), are more susceptible to the capsid binding compound pleconaril. However, despite having an atypical branched pocket factor and a seemingly incompatible hydrophobic pocket, CV-A24v is robustly inhibited by pleconaril. Binding of pleconaril is permitted by selection of a non-precluding conformation of Tyr\textsuperscript{161} which allows the drug molecule to bind deeper in the hydrophobic pocket compared to the natural pocket factor. Pleconaril has previously been investigated as a potential treatment for rhinovirus infections and has undergone clinical trials towards this end. When administered orally, pleconaril was found to reduced disease severity and duration in adults but ultimately was not licensed, primarily because it interfered with birth control measures, thus increasing the risk of unplanned pregnancy (Hayden et al., 2003). A subsequent study, in which pleconaril was administered using a nasal spray, failed to show a statistically-significant clinical benefit (Merck Sharp & Dohme Corp, 2006). On the other hand, pleconaril has shown promise as a potential treatment of enterovirus related neonatal sepsis, having
been shown to improve survival rates in this disease setting when administered orally (Abzug et al., 2016). To date, pleconaril has not been considered as a treatment for enteroviruses with an ocular tropism. In light of the results presented here, pleconaril could be considered a potential treatment for AHC caused by CV-A24v. One could envisage that pleconaril could be administered as a topical treatment, similarly to chloramphenicol (Silvester et al., 2016) and trifluridine (Mallika et al., 2008) which can be given as eye drops to treat bacterial and herpes/vaccinia eye infections, respectively. Since CV-A24v is believed to be transmitted via hand-to-eye contact, the fate of CV-A24v after the initial disposition into the eye might be largely determined by its capacity to attach to the conjunctival tissue before it is cleared by the continuous draining of tear film through the nasolacrimal duct. Hypothetically, topical administration of pleconaril could ensure a high local concentration of the drug molecule in the eye and lacrimal drainage system. It has previously been suggested that inhibitors of sialic acid binding could be used as a potential topical treatment for CV-A24v (E.C. Nilsson et al., 2008; Zocher et al., 2014). Such inhibitors could be used in conjunction with pleconaril as a combination therapy. Going forward, studies into the safety, feasibility and efficacy of using pleconaril to treat AHC should be performed. Furthermore, the susceptibility of CV-A24v to other capsid-binding compounds could be determine, such as GPP3, NLD and ALD (De Colibus et al., 2015; De Colibus et al., 2014; Kelly et al., 2015). Furthermore, the 1.8 Å resolution crystal structure of CV-A24v in complex with pleconaril provides a detailed platform for the structure-based design of novel capsid binding compounds with improved activity against this pandemic pathogen.
Chapter 5 – Summary, future perspectives and concluding remarks

5.1 Summary

This thesis describes high-resolution cryo-EM structures of two important human pathogens and their receptor complexes. These studies provide insights into the early events in the lifecycle of these viruses, and more generally, they demonstrate the utility of cryo-EM for studying the interactions of non-enveloped viruses with glycan and protein receptors at better than 4 Å resolution.

Chapter one provides a brief historical account of the structural techniques used to study viruses, namely X-ray crystallography and cryo-EM, and provides an overview of the advantages and disadvantages of both these methods. Furthermore, the different aspects of viruses which can be structurally studied are described, including genome packaging and receptor interactions.

Chapter three describes structural studies of BKV, starting with intermediate resolution reconstructions of the virion and a VP1 only VLP. These sub-nanometre reconstructions suggest a role of the minor capsid proteins in genome recognition or condensation and provide clues as to how the genome may be packaged in the boundary layer beneath the capsid surface. The second part of chapter three described the high-resolution structure of BKV, the first for an infectious human polyomavirus. This structure highlights differences in the capsid compared to non-human polyomaviruses studied previously. Furthermore, the complex of BKV and a fragment of its largest sialylated receptor, GT1b, is described, which suggests a mechanism of how this branched molecule enhances infectivity over the smaller ganglioside variants studied previously. Finally, the structure of BKV in complex with the model GAG heparin is presented, which identifies a positively charged intercapsomere region as a candidate binding site.

Chapter four describes results from a collaborative project involving a number of institutions across Europe. ICAM-1 is identified as an essential protein
receptor for CV-A24v and using cryo-EM, the molecular basis of this virus-receptor interaction is established. Also, it is shown that the presence of ICAM-1 is associated with an increased number of empty virions, thus suggesting that it acts as the uncoating receptor for CV-A24v. By determining the 3.8 Å resolution structure of the empty CV-A24v particle, the conformational changes associated with particle expansion and genome release are revealed. While the uncoating mechanism of CV-A24v is similar to that described for other enteroviruses; the hydrophobic pocket has an atypical configuration which appears as if it would preclude binding of capsid binding compounds. Despite this, CV-A24v is shown to be sensitive to the capsid binding compound pleconaril and using X-ray crystallography, the structural changes required to accommodate this drug molecule is revealed.

5.2 Continuation of the projects and future perspectives

5.2.1 Studies on BKV

One logical next step in the BKV project described would be to perform mutagenesis studies on the cationic residues surrounding the putative site of heparin binding. Such an approach could utilise sialic acid binding deficient pseudo-viruses, as described previously (Geoghegan et al., 2017). Furthermore, using low-molecular weight heparin with a minimum saccharide length of ~eight units may improve the resolution of the bound GAG molecule and enhance interpretability of the binding site. Alternatively, it may be possible to use 3D classification methods to separate the presumably heterogeneous mixture of binding modes which are present. As stated in chapter three, it would be interesting to determine complexes of BKV with antibodies to determine the molecular basis of neutralisation. Indeed, a currently unpublished cryo-EM structure of a BKV VLP in complex with a single-chain neutralising antibody, at 4.2 Å resolution, is listed on the EMDB under the accession code EMD-4398. Furthermore, the cryo-EM workflow described for BKV could be applied a number of clinically important human polyomavirus such as JCV and MCV.

Longer term investigations of BKV could employ an in situ cryo-ET approach (Asano et al., 2016). Whilst technically very challenging, such experiments
could reveal new biological insights into the process of cell entry, trafficking, uncoating, assembly and release. It would also be interesting to further investigate the organisation of the viral minichromosome with the virus particle, assuming there is a unique structure. Such studies would undoubtedly utilise computational methods to resolve the symmetry mismatch between the genome and icosahedral capsid, as described for transcribing and non-transcribing cypoviruses (Liu and Cheng, 2015). Single-particle EM studies could also be performed on individual pentons in complex with nucleosomes to further investigate the structure and role of the minor capsid proteins.

5.2.2 Studies on CV-A24v

Given that CV-A24v is susceptible to pleconaril, it would be fascinating to understand the susceptibility of CV-A24v to other know capsid binding compounds such as GPP3, NLD and ALD, as described for EV-A71 and CV-A16 (De Colibus et al., 2014; De Colibus et al., 2015). Furthermore, the 1.8 Å resolution crystal structure of CV-A24v in complex with pleconaril could be used for in silico design of more potent inhibitors of this virus. Given that CV-A24v has spatially distinct sites for ICAM-1 and sialic acid binding, it would be interesting to determine its complexes with neutralising antibodies to see which structural features of the virus are the dominant epitopes. Such studies could perhaps pave the way for the development of bivalent antibody compositions. More generally, the on-grid binding approach used to capture the complex between CV-A24v and ICAM-1 could be applied to a number of enteroviruses whose receptor interactions have not yet been structurally characterised.

It is now well established that icosahedral averaging of virus particles can obscure interesting structural features. There have been great strides in the development of computational and biochemical methods to capture asymmetric states of viruses. One such approach is to use lipid nanodiscs containing virus receptors, in order to stimulate only part of the virus capsid, as one would expect in a natural infection. This methodology was employed to study the asymmetric entry intermediates of CV-B3 (Lee et al., 2016) and poliovirus (Nasr et al., 2017), and a similar approach could be employed for
CV-A24v. In addition, as stated above for BKV, *in situ* cryo-ET studies could be used to further examine aspects of the CV-A24v lifecycle.

### 5.3 Concluding remarks

Whilst the pathology of BKV and CV-A24v differ greatly, both are clinically important human pathogens responsibly for debilitating diseases which currently lack targeted treatments. Collectively the structures described in this thesis provide insights into capsid structure, genome packaging, receptor binding, genome uncoating and inhibition of these viruses. More generally, a simple and robust methodology for on-grid binding of receptors is described which may prove useful in determining the structures of currently uncharacterised virus-receptor interactions. The structures of BKV-GT1b and CV-A24v-ICAM-1 presented in this thesis are the first virus cryo-EM structures to resolve bound glycan and protein receptors at better than 4 Å resolution. At this resolution, side chains become visible and the interpretability of these important biological interactions is greatly improved. Such results demonstrate the utility of cryo-EM to study these interactions at high-resolution. Indeed, unpublished sub-4 Å resolution virus-receptor complexes are poised to be released, including Seneca valley virus in complex with anthrax toxin receptor 1 (EMD-7772) and EV-D68 in complex with 6'-sialyl-N-acetyllactosamine (EMD-7638). This is testament to fast pace that the field of cryo-EM is advancing, driven by improvements in sample preparation, data collection and computational tools used to generate 3D reconstructions. The ever-improving resolution of single-particle approaches ensures this technique will continue to make huge contributions to knowledge of virus structure. Furthermore, technological improvements and increased accessibility of cryo-ET will likely facilitate a move towards studying aspects of virus lifecycles within the infected cell.
List of References


Academy of Sciences of the United States of America. 100(18), pp. 10477–10482.


the National Academy of Sciences of the United States of America. 113(41), pp. 11519–11524.


from reversible exposure of these sequences at physiological temperature. *Journal of Virology*. **68**(6),pp.3965–3970.


Target Viral Protein 2C. Antimicrobial agents and chemotherapy. 60(5),pp.2627–2638.


wise design of inhibitors to 3C proteinase from enterovirus 71. *Biochimica et Biophysica Acta (BBA) - General Subjects*. **1860**(6),pp.1299–1307.


Yoder, J.D., Cifuente, J.O., Pan, J., Bergelson, J.M. and Hafenstein, S. 2012. The crystal structure of a coxsackievirus B3-RD variant and a refined 9-angstrom cryo-electron microscopy reconstruction of the virus complexed with decay-accelerating factor (DAF) provide a new footprint of DAF on the virus surface. Journal of Virology. 86(23), pp.12571–12581.


Appendix A  Additional publications

During the course of this PhD project, the candidate made contributions to additional co-authored publications, the details of which are included below:


Abstract

Histidine biosynthesis is an essential process in plants and microorganisms, making it an attractive target for the development of herbicides and antibacterial agents. Imidazoleglycerol-phosphate dehydratase (IGPD), a key enzyme within this pathway, has been biochemically characterised in both Saccharomyces cerevisiae (Sc_IGPD) and Arabidopsis thaliana (At_IGPD). The plant enzyme, having been the focus of in-depth structural analysis as part of an inhibitor development program, has revealed details about the reaction mechanism of IGPD, whereas the yeast enzyme has proven intractable to crystallography studies. The structure–activity relationship of potent triazole-phosphonate inhibitors of IGPD has been determined in both homologs, revealing that the lead inhibitor (C348) is an order of magnitude more potent against Sc_IGPD than At_IGPD; however, the molecular basis of this difference has not been established. Here we have used single-particle electron microscopy (EM) to study structural differences between the At and Sc_IGPD homologs, which could influence the difference in inhibitor potency. The resulting EM maps at ∼3 Å are sufficient to de novo build the protein structure and identify the inhibitor binding site, which has been validated against the crystal structure of the At_IGPD/C348 complex. The structure of Sc_IGPD reveals that a 24-amino acid insertion forms an extended loop region on the enzyme surface that lies adjacent to the active site, forming interactions with the substrate/inhibitor binding loop that may influence inhibitor potency. Overall, this study provides insights into the IGPD family and demonstrates the power of using an EM approach to study inhibitor binding.
Contribution

*Joint first author. The candidate processed, built and refined the At\_IGPD cryo-EM structure, generated most of the figures and contributed towards writing the first draft.


Abstract

BK polyomavirus (BKPyV; hereafter referred to as BK) causes a lifelong chronic infection and is associated with debilitating disease in kidney transplant recipients. Despite its importance, aspects of the virus life cycle remain poorly understood. In addition to the structural proteins, the late region of the BK genome encodes for an auxiliary protein called agnoprotein. Studies on other polyomavirus agnoproteins have suggested that the protein may contribute to virion infectivity. Here, we demonstrate an essential role for agnoprotein in BK virus release. Viruses lacking agnoprotein fail to release from host cells and do not propagate to wild-type levels. Despite this, agnoprotein is not essential for virion infectivity or morphogenesis. Instead, agnoprotein expression correlates with nuclear egress of BK virions. We demonstrate that the agnoprotein binding partner α-soluble N-ethylmaleimide sensitive fusion (NSF) attachment protein (α-SNAP) is necessary for BK virion release, and siRNA knockdown of α-SNAP prevents nuclear release of wild-type BK virions. These data highlight a novel role for agnoprotein and begin to reveal the mechanism by which polyomaviruses leave an infected cell.

Contribution

The candidate propagated and purified wild-type and mutant BKV, characterised virion morphology by negative stain electron microscopy and wrote the corresponding section of the manuscript.
Abstract

Cryo-electron microscopy (cryo-EM) can now be used to determine high-resolution structural information on a diverse range of biological specimens. Recent advances have been driven primarily by developments in microscopes and detectors, and through advances in image-processing software. However, for many single-particle cryo-EM projects, major bottlenecks currently remain at the sample-preparation stage; obtaining cryo-EM grids of sufficient quality for high-resolution single-particle analysis can require the careful optimisation of many variables. Common hurdles to overcome include problems associated with the sample itself (buffer components, labile complexes), sample distribution (obtaining the correct concentration, affinity for the support film), preferred orientation, and poor reproducibility of the grid-making process within and between batches. This review outlines a number of methodologies used within the electron-microscopy community to address these challenges, providing a range of approaches which may aid in obtaining optimal grids for high-resolution data collection.

Contribution

The candidate co-wrote the section on practical applications of continuous carbon, provided figures for two sections and assisted in writing, reviewing and editing of the manuscript.