An Analysis of Human Cytomegalovirus Gene Family Function

A thesis submitted in candidature for the degree of

DOCTOR OF PHILOSOPHY

by

Hester Amy Nichols

September 2018

Division of Infection and Immunity,

School of Medicine,

Cardiff University

Declarations

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

Signed…... (candidate)   Date................................

STATEMENT 1

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

Signed…... (candidate)   Date................................

STATEMENT 2

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

Signed…... (candidate)   Date................................

STATEMENT 3

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

Signed…... (candidate)   Date................................

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

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

Signed…... (candidate)   Date................................

Acknowledgements

I would firstly like to express my gratitude to my supervisors, Dr Ceri Fielding and Prof. Gavin Wilkinson for giving me the opportunity to do a PhD, and for their support, guidance, valued insights and patience throughout, alongside their support of me attending conferences to further my professional and personal development. I would also like to thank my supervisor Dr Peter Tomasec who sadly passed away last year.

I am also grateful to Dawn Roberts for her technical support, and would like to thank everyone within the Wilkinson lab, for their help in the lab and their valuable input within team meetings. I would especially like to thank everyone in the lab and the PhD students in the building for their friendship over the last 3 years, and for making the lab such an enjoyable workplace. I would especially like to thank Simone Forbes for her constant enthusiasm and encouragement throughout my writing up period; and Carmen Bedford who I was lucky enough to meet at the start of my Cardiff adventure and continue my PhD journey with, and who has kept me sane with multiple cups of tea and chick flicks over the years.

I would also like to thank my funding body, the Medical Research Council (MRC), and our collaborators in Glasgow (in the laboratory of Dr Andrew Davison, MRC-University of Glasgow Centre for Virus Research) and in Cambridge (in the laboratory of Dr Paul Lehner, Cambridge Institute for Medical Research) for their help with my project.

Finally, I would like to thank my boyfriend Jason Vincent and my family, especially my mum, dad and sister, for supporting me and believing in me throughout my PhD studies. I couldn't have gotten through it without you all!

Summary

Human cytomegalovirus (HCMV) is a β-herpesvirus that causes complications in immuno-compromised individuals and is the leading infectious cause of birth defects. The HCMV genome contains 15 gene families, which contain between 2 to 14 members. One of these, the US12 gene family, consists of a sequential cluster of 10 genes (*US12* to *US21*) that are highly conserved in clinical isolates. This family has roles in tropism and immune evasion and was recently found to regulate the cell surface expression of a wide array of immune ligands. This included the regulation of ligands for the natural killer (NK) cell activating receptors NKG2D and NKp30 (MICA and B7-H6 respectively), which were targeted by US18 and US20. To complement these mechanistic studies, a C-terminal V5 epitope tag was added to each US12 family gene within the HCMV Merlin genome. A large proportion of the US12 family were shown to be degraded within the cell, possibly within lysosomes, which suggests that they may interact with their targets proteins in order to redirect them for degradation. Expression of US12 family members was detectable by immunoblotting during an infection time-course, with many US12 family members expressed during the Tp3 temporal class of HCMV gene expression. Three members of the family were also demonstrated to be N-glycosylated during HCMV infection. The US12 family appear to have associations with the virion assembly compartment, and correspondingly, 7 US12 family members are found within the virion. Furthermore, the majority of the US12 family also show co-localisation with endoplasmic reticulum-derived membranes. These data build on our previous functional characterisation to give insights into the workings of this important HCMV gene family.

Table of Contents

List of Figures

List of Tables

Abbreviations

°C	degrees centigrade
α	alpha
β	beta
γ	gamma
μg	microgram
μl	microlitre
Ad	adenovirus
ADCC	antibody dependent cell-mediated cytotoxicity
AIDS	acquired immune deficiency syndrome
APC	antigen-presenting cell
BAC	bacterial artificial chromosome
BLAST	basic local alignment search tool
bp	base pair
BoHV	bovine herpesvirus
BSA	bovine serum albumin
CCMV	chimpanzee cytomegalovirus
CCV	channel catfish virus
ChHV	chelonid herpesvirus
CMV	cytomegalovirus
CTL	cytotoxic T lymphocyte
CX3CR1	CX3C chemokine receptor 1
DAMPS	damage-associated molecular patterns
DAPI	4, 6' diamino-2-phenylindole
DB	dense body
DCs	dendritic cells
DMEM	Dulbecco's modified Eagle's medium
DMS	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates

DNase	deoxyribonuclease
DOG	2-deoxygalactose
dsDNA	double stranded DNA
DTT	dithiothreitol
DURP	dUTPase-related protein
E	early
EBV	Epstein Barr virus
ECs	endothelial cells
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EIHV-1	elephantid herpesvirus 1
EGFR	epidermal growth factor receptor
EHV	equid herpesvirus
EM	electron microscopy
Endo H	endoglycosidase H
ER	endoplasmic reticulum
ERGIC	ER-Golgi intermediate compartment
ESCRT	endosomal sorting complexes required for transport
EHV	equine herpesvirus
FBS	foetal bovine serum
g	grams
GaHV	avian (Gallid) herpesvirus 2
GalK	galaktokinase
gB	glycoprotein B
gH	glycoprotein H
gL	glycoprotein L
gM	glycoprotein M
GMCMV	green monkey CMV
gN	glycoprotein N
gO	glycoprotein O
gp	glycoprotein

GPCMV	guinea pig CMV
GPCR	G-protein couple receptor
GvHD	graft-versus-host disease
HAART	highly active antiretroviral therapy
HCMV	human cytomegalovirus
HCT	haemopoietic (stem) cell transplant
HDAC	histone deacetylase
hDAXX	human death-domain associated protein
HHV	human herpesvirus
HF-CARS	human fibroblasts with human Coxsackie adenovirus receptor
HF-TERTS	human fibroblasts immortalised with telomerase reverse transcriptase
HFFs	human foetal foreskin fibroblasts
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
HLFs	human lung fibroblasts
HMVEC	human microvascular endothelial cell
HRP	horseradish peroxidise
HSV	herpes simplex virus
hpi	hours post infection
HVA	Herpesvirus ateles (Spider monkey herpesvirus)
ICTV	International Committee on the Taxonomy of Viruses
IE	immediate early
IFN	interferon
Ig	immunoglobulin
IL	interleukin
IV	intravenous
INM	inner nuclear membrane
IPTG	isopropyl β-D-1-thiogalactopyranoside
IR	internal repeat
IRF	interferon regulatory factor
IRS	internal repeat short

IRL	internal repeat long
ITAMs	immuno-receptor tyrosine-based activating motifs
ITIMs	immuno-receptor tyrosine-based inhibitory motifs
Kb	kilobase
Kbp	kilobase pairs
kDa	kilodalton
KIRs	killer immunoglobulin-like receptors
KSHV	Kaposi's sarcoma-associated herpesvirus
l	litres
L	late
LA	latency-associated
LB	Luria-Bertani
LIRs	leukocyte immunoglobulin-like receptors
lncRNAs	long non-coding RNAs
LUNA	latent undefined nuclear antigen
MAMPS	microbe-associated molecular patterns
MAPK	mitogen-activated protein kinase
MBCs	memory B cells
MCMV/MuHV	mouse cytomegalovirus/ murid herpesvirus
MCP	major capsid protein (pUL86)
mCP	minor capsid protein (pUL46)
mC-BP	minor capsid protein-binding protein (pUL85)
MHC	major histocompatibility complex
MICA	major histocompatibility complex class I-related chain A
MICB	major histocompatibility complex class I-related chain B
MIE	major immediate early
MIEP	major immediate early promoter
MIP-1α	macrophage inflammatory protein 1-alpha (CCL3)
miRNAs	micro ribonucleic acids
MOI	multiplicity of infection
mRNA	messenger RNA

MSHV	*Miniopterus schreibersii* herpesvirus
MTOC	microtubule-organizing centre
MVA	modified vaccinia ankara vector
NCR	natural cytotoxicity receptor
NEC	nuclear egress complex
NIEP	non-infectious enveloped particles
NF-κB	nuclear factor kappa B (NF kappa-light-chain-enhancer of activated B cells)
NK cell	natural killer cell
NKG2D	natural killer group 2, member D
NLRs	NOD-like receptors
nm	nano-metre
NOD	nucleotide-binding oligomerization domain
NOS	reactive nitrogen species
NP40	Nonyl phenoxypolyethoxylethanol-40
NPC	nuclear pore complex
nt	nucleotide
NuRD	nucleosome remodelling and deacetylase
OD	optical density
OMCMV	owl monkey CMV
ONM	outer nuclear membrane
ORF	open reading frame
oriLyt	lytic origin of DNA replication
PAMPs	pathogen-associated molecular patterns
pAP	precursor assembly protein (pUL80.5)
PBMC	peripheral blood mononuclear cell
PBS	phosphate-buffered saline
PCMV	porcine cytomegalovirus
PCR	polymerase chain reaction
pDCs	plasmacytoid dendritic cells
PDGFR	platelet-derived growth factor receptor
pfu	plaque-forming unit

PI3K	phosphatidylinositol-3 kinase
PKR	protein kinase R
PM	plasma membrane
PML	promyelocytic leukemia
pNP1	proteinase precursor (pUL80A)
PORT	portal protein (pUL104)
pp	phosphoprotein
PNGase F	peptide N-glycosidase F
pp	phosphoprotein
RANTES	regulated on activation, normal T cell expressed and secreted chemokine
RCMV	rat cytomegalovirus
RCMVE	rat cytomegalovirus England
RhCMV	rhesus cytomegalovirus
RNA	ribonucleic acid
RNP	ribonucleoprotein
ROS	reactive oxygen species
RPE	retinal pigment epithelial
SCP	small capsid protein (pUL48A)
SILAC	stable isotope labelling by amino acids in cell culture
SLAM	signalling lymphocyte-activation molecule
SMCMV	squirrel monkey CMV
TAE	tris-acetate-EDTA
TAP	transporter associated with antigen processing
TBST	tris-buffered saline Tween-20
TCRs	T cell receptors
TGN	trans-Golgi network
THV	tupaiid herpesvirus
TLR	toll-like receptor
TNF	tumour necrosis factor
TRAIL	TNF-related apoptosis-inducing ligand
TR	terminal repeat

TRL	terminal repeat long
Tp	temporal kinetics class
TRS	terminal repeat short
UL	unique long
UL/*b'*	unique long b prime region
ULBP	UL16-binding protein
US	unique short
UV	ultraviolet
vDNA	viral DNA
VEEV	Venezuelan equine encephalitis virus
vICA	viral inhibitor of caspase-8 activation
vMIA	viral mitochondrion-localised inhibitor of apoptosis
VZV	varicella-zoster virus
vAC	virion assembly compartment
WCL	whole cell lysate
X-gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside

1 Introduction

1.1 Herpesviruses

Herpesviruses (or *Herpesviridae*) are a large group of DNA viruses that infect a wide range of species including primates, birds, amphibians, reptiles and invertebrates. Over 200 have been identified to date, each of which is usually restricted to one species alone although hosts can be infected with more than 1 herpesviruses at a time (Fields et al., 2013). Due to the size of the group, they have recently been classified into 3 further groups by the International Committee on the Taxonomy of Viruses (ICTV, 2011) into *Herpesviridae* (herpesviruses of mammals, birds and reptiles), *Alloherpesviridae* (herpesviruses of fish and amphibians) and Malacoherpesviridae (herpesviruses of bivalves). Herpesviruses are also split into three mammalian subfamilies- alpha (α), beta (β) and gamma (γ) – that arose around 180-220 million years ago. The emergence of divergence of species within sub-lineages has expanded rapidly over the last 80 million years, most likely prior to or alongside speciation of the host lineages (McGeoch et al., 1995). Each sub-family contain multiple genera of viruses (**Table 1.1**). α-herpesviruses generally have short replication cycles and have the most variable host range of these 3 groups, infecting bird and reptile classes as well as mammals (Whitley, 1996, Pellett and Roizman, 2013). Members of the α-herpesvirus sub-family can spread rapidly in cell culture, set up latent infections primarily in sensory ganglia, but have a wide host range. Herpes simplex virus (HSV)-1 and HSV-2 for example can infect neuronal cells, leukocytes, epithelial cells and fibroblasts (Spear and Longnecker, 2003) and rapidly destroy the cells they infect. β-herpesviruses are mid-range in that their host range is restricted, but generally less so than the γ-herpesviruses, and tend to have long reproductive cycles, which can be over 7 days. Their infection in cultured cells is relatively slow and infected cells frequently become enlarged as demonstrated by cytomegaloviruses (Whitley, 1996). γ-herpesviruses on the other hand tend to have a more restricted host range within the family or order of the natural host (Pellett and Roizman, 2013). This sub-family are usually specific for lymphoid cells but can also infect other cells types, e.g. Epstein Barr virus (EBV) persistence is associated with B lymphocytes yet the virus also replicates in epithelial cells (Grinde, 2013).

Herpesviruses have evolved alongside (and are generally well adapted to) their hosts, causing life-long infection, with all known members to date having the ability to both remain latent (or in a low replicative state) in their natural host, and with the ability to re-activate (Pellett and Roizman, 2013, Grinde, 2013). They all have a characteristic spherical virion morphology consisting of an envelope, tegument, capsid and a core. All

Table 1.1: Current mammalian sub-groups of herpesviruses and their genera, adapted from Davison (2007b) and updated using data from the ICTV*

Sub-group	Genera◊	Examples (non-exclusive)
Alpha	Simplexvirus	Herpes Simplex virus types 1 and 2 (HSV-1 and HSV-2), Spider monkey herpesvirus (Herpesvirus ateles 1, HVA-1)
	Varicellovirus	Varicella-Zoster virus (VZV), Equine herpesvirus (EHV)
	Mardivirus	Avian (Gallid) herpesvirus 2 (GaHV-2)
	Iltovirus	Avian (Gallid) herpesvirus 1 (GaHV-1)
	Scutavirus	Chelonid herpesvirus 5 (ChHV-5)
Beta	Cytomegalovirus (CMV)	Human CMV (HHV-5/HCMV)
	Muromegalovirus	Murid herpesviruses/Mouse cytomegaloviruses 1, 2 and 8 (MuHV-/MCMV- 1,2 and 8)
	Roseolovirus	Human herpesviruses (HHV-) 6A, 6B and 7
	Proboscivirus	Elephantid herpesvirus 1 (EIHV-1)
Gamma	Lymphocryptovirus	Human herpesvirus 4/Epstein-Barr virus (HHV-4/EBV)
	Rhadinovirus	Bovine herpesvirus 4 (BoHV-4), Kaposi's Sarcoma-associated herpesvirus (KSVH)
	Macavirus†	Bovine herpesvirus 6 (BoHV-6)
	Percavirus†	Equid herpesvirus 2 and 5 (EHV-2,5)
Undefined	Ictalurivirus ✿	Ictalurid herpesvirus 1/Channel catfish virus (CCV)

† More recently defined genera
◊ Unassigned genera within each group (1 unassigned genera in alpha and beta, and two in the gamma sub-family) have not been included
✿ Ictalurivirus is an additional genus which has not yet been assigned to a sub-family

herpesviruses have large linear, double-stranded DNA genomes (125-240 kbp) with viral DNA replication and capsids assembly completed in the nucleus (Whitley, 1996, Davison, 2007b). Their DNA is contained in the core, which is encapsulated by the icosahedral capsid. The nucleocapsid is then surrounded by tegument (which has a poorly defined structure) and enclosed by a lipid envelope that generally contains ten or more viral membrane glycoproteins, some of which are relatively conserved across different family members and some of which vary widely (Davison, 2007b). A common feature of herpesviruses is that they encode their own enzymes for viral nucleic acid biosynthesis and that productive replication and virion release tends to be linked to cell death (Whitley, 1996). Their structural similarity suggests a common ancestor and, although their classification was originally based solely upon this virion structure (Fields et al., 2013), newer members tend to be classified primarily on the basis of their genomic sequence. Herpesviruses are designated by the species they infect, their family or sub-family, and then by sequential numbering, such as 'Human herpesvirus 7' (HHV-7). These viruses often have a common name as well, such as 'human herpesvirus 3' which is known as 'varicella-zoster virus' (or VZV).

One commonality of herpesvirus genomes is the conservation of their sequence organisation, and they can therefore be grouped into classes based on the copy number, location, and orientation of repeat elements. (Roizman and Pellett, 2001, Barry and Chang, 2007). These 6 structure classes (named A-F) correspond to different genome layouts (**Figure 1.1**). Class A genomes contain a unique sequence flanked by a direct repeat, and is represented by viruses such as HHV-6 and HHV-7 (Davison, 2007a). Class B defines genomes that have direct repeated sequences at the termini that consist of variable copy numbers of a tandemly repeated sequences, and contains mostly γ-herpesviruses, such as HHV-8. The class C structure is a derivative of class B, where the internal direct repeats are unrelated to the terminal set, and this structure is found in viruses such as EBV (Davison, 2007a). Class D genomes are characteristic of alpha-herpesviruses in the Varicellovirus genus, such as VZV, and contain two unique regions each flanked by inverted repeats (Davison, 2007a). Class E genomes are the most complex structure type and are generally characteristic of *Simplexviruses* (α-herpesviruses) but have also evolved independently in the β-herpesvirus lineage in the cytomegaloviruses (CMVs) of higher primates such as humans and chimpanzees (HCMV and CCMV respectively) (Weststrate et al., 1980, Davison et al., 2003b). Class E genomes are similar to class D, as they also contain internal unique sequences whereby both termini are repeated
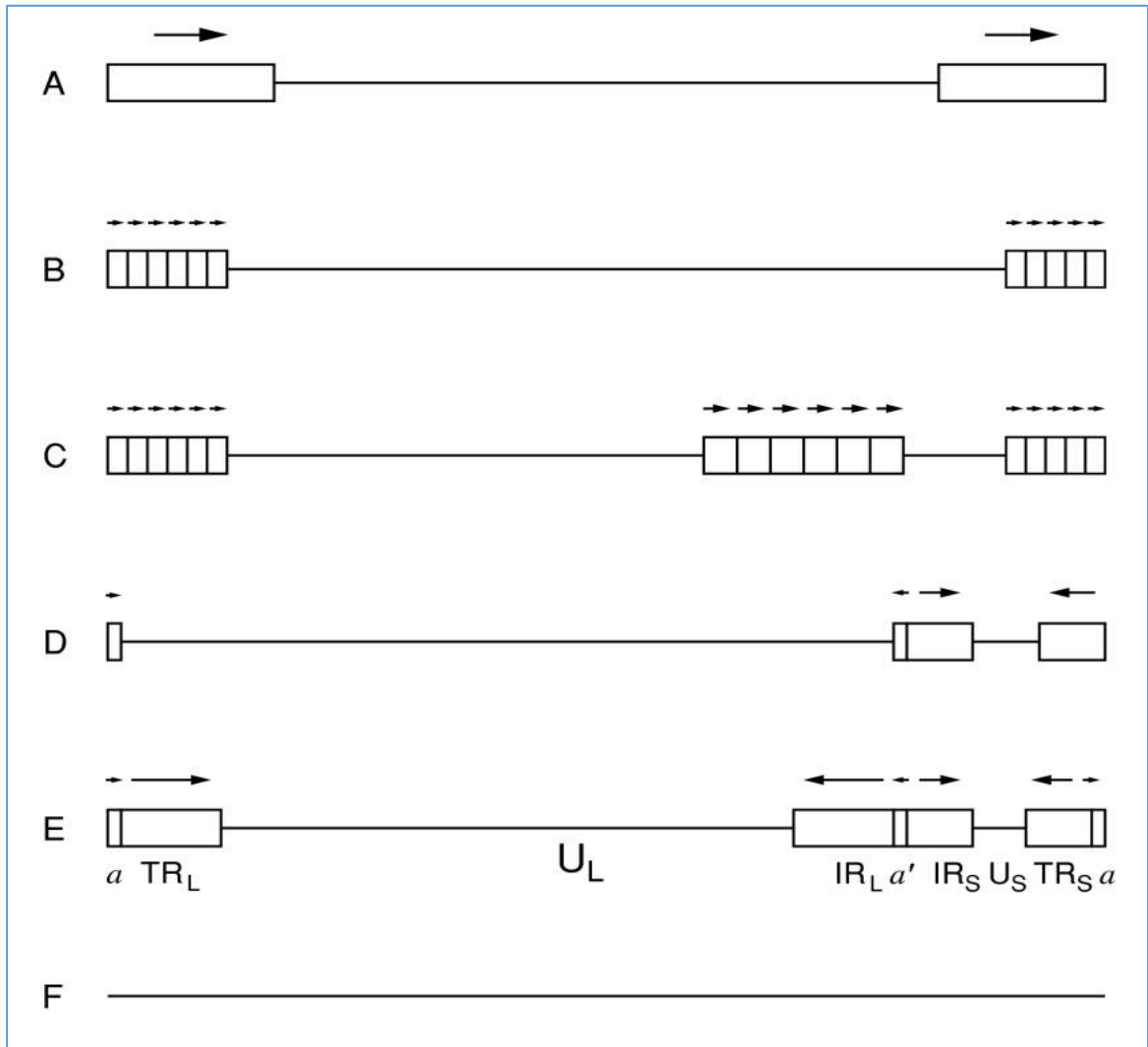
Figure 1.1: The 6 classes of herpesvirus genome structure. Genome structure classes A-F figure taken from A. J. Davison (2007a) as initially described by (Roizman & Pellett, 2001). Horizontal lines depict the unique regions and rectangles depict the repeat regions of the genome, with orientations shown with arrows. Regions are detailed for class E only; with regions named as follows, US= Unique short, UL= Unique long, TR/IR= Terminal/internal repeat (L, Long or S, Short). a denotes terminal redundancy and a' the internal inverted copy. Genomes are not to scale.

in an inverted orientation, except that the repeat regions are much larger and segment inversion gives rise to four equimolar genome isomers (Davison, 2007a, Whitley, 1996). These unique regions are termed unique long (UL) and unique short (US). Class F genomes lack the inverted and direct repeats across other herpesvirus genomes, and is represented by a beta-herpesvirus Tupaiid herpesvirus (THV), but not by any human viruses (Davison, 2007a).

The core genes conserved down through herpesvirus evolution are found in seven blocks near the centre of the genome, encoding proteins involved in nucleic acid metabolism, DNA replication, virion structure and maturation (Barry and Chang, 2007, Roizman and Pellett, 2001). This includes the conservation of three gene families (UL25, UL82, and US22) (Barry and Chang, 2007). Across β-herpesvirus genomes, there are also three blocks of conserved genes, and a G-protein coupled receptor. This has now also been expanded to include blocks *UL23-43*, *UL82-84*, and *US22-26*, along with two loci whose location, structures and splicing patterns are highly preserved (Barry and Chang, 2007, Davison et al., 2003b). The β-herpesvirus sub-family cluster into 4 distinct groups as previously mentioned, as seen by phylogenetic tree analysis of core genes (**Figure 1.2 A**). All 4 groups however share a common appearance, prolonged replication cycles, species specificity, and their tropism for the salivary gland (Mocarski et al., 2013).

Across the range of sub-families, there are 9 herpesviruses in total that can infect humans as their primary hosts (**Table 1.2**). Herpes simplex viruses 1 and 2 along with varicella-zoster virus are α-herpesviruses; human cytomegalovirus (HCMV), human herpesviruses 6A, 6B and 7 (HHV-6A, 6B and 7) are β-herpesviruses; and Kaposi's sarcoma-associated herpesvirus (HHV-8/KSVH) and Epstein-Barr virus (EBV) are γ-herpesviruses (Pellett and Roizman, 2013). These human viruses and their relationships are demonstrated in **Figure 1.2B,** with their grouping into α, β and γ sub-groups made apparent**.**

1.2 Human Cytomegalovirus (HCMV)

The cytomegalovirus genus is one of the four β-herpesvirus genera that can infect a range of primate species including humans (HCMV), chimpanzees (CCMV), African green monkeys (SCMV), New World monkeys (Aotine herpesviruses 1 and 3) and Rhesus monkeys (RhCMV). CMVs are known as salivary gland viruses and have common growth characteristics with nuclear and cytoplasmic inclusions contributing to their characteristic
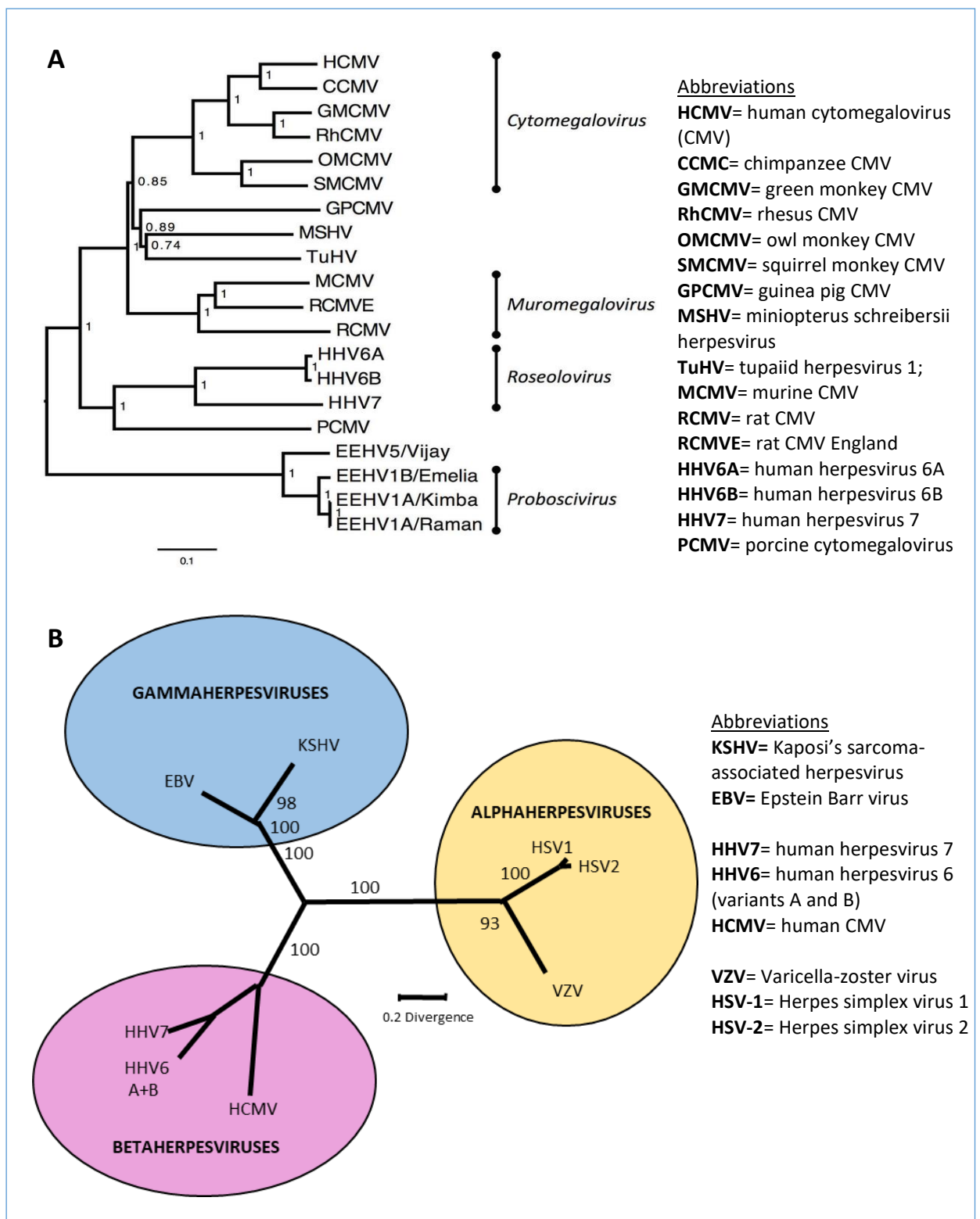
Figure 1.2: Phylogenetic analysis of herpesviruses. **A)** Phylogenetic tree of the 4 β-herpesvirus genera using amino acid sequences of core genes (U38, U39, U40, U41, U57, U60, U77, and U81) as created by Wilkie et al. (2014). The scale bar shows nucleotide differences/nucleotide. **B)** Phylogenetic tree of the human herpesviruses into their sub-families, edited from Moore et al. (1996), based on aligned amino acid sequences between for the Major capsid protein (MCP). Branch lengths are based on evolutionary distance.

Table 1.2: The 9 Herpesviruses that infect humans as their primary host, as modified from Pellett and Roizman (2013)◊

Virus designation	Common virus name	Sub-family (genus)	Genome size
Human Herpesvirus 1 (HHV-1)	Herpes simplex virus 1 (HSV-1)	α-herpesvirus (Simplexvirus)	152 Kbp
Human Herpesvirus 2 (HHV-2)	Herpes simplex virus 2 (HSV-2)	α-herpesvirus (Simplexvirus)	152 Kbp
Human Herpesvirus 3 (HHV-3)	Varicella-zoster virus (VZV)	α-herpesvirus (Varicellovirus)	125 Kbp
Human Herpesvirus 4 (HHV-4)	Epstein Barr virus (EBV)	γ-herpesvirus (Lymphocryptovirus)	172 Kbp
Human Herpesvirus 5 (HHV-5)	Human Cytomegalovirus (HCMV)	β-herpesvirus (Cytomegalovirus)	Merlin 236 Kbp, AD169 230 Kbp†
Human Herpesvirus 6 (HHV-6)	HHV-6 A	β-herpesvirus (Roseolovirus)	159/170* Kbp
	HHV-6 B	β-herpesvirus (Roseolovirus)	162/168* Kbp
Human Herpesvirus 7 (HHV-7)	HHV-7	β-herpesvirus (Roseolovirus)	145 Kbp
Human Herpesvirus 8 (HHV-8)	Kaposi's sarcoma-associated herpesvirus (KSVH)	γ-herpesvirus (Rhadinovirus)	170/210* Kbp

◊ collated from the data provided by the Herpesvirales Study Group of the International Committee on Taxonomy of Viruses (ICTV, 2011).
* Values obtained in different labs
† 2 well-known strains of HCMV given as examples

cytopathology which were first noted in 1881 but was originally thought to be caused by protozoa (Ribbert, 1904). They were named due to this common enlargement of the infected cells, as 'cytomegalovirus' (cytomegalia meaning large cell) and they were subsequently linked to 'generalized cytomegalic inclusion disease' seen in children (Wyatt et al., 1950, Rowe et al., 1953). Viruses were historically hard to culture, and it wasn't until after 1949 when the successful cultivation and isolation of poliomyelitis virus was carried out using human embryonic cells, that HCMV was isolated in a cell culture system (Enders et al., 1949, Rowe et al., 1953).HCMV, also designated human herpesvirus 5 (HHV-5), can now be cultured in a range of human cells (**Section 1.4.1.1**). HCMV has long been perceived as a slowly replicating virus based on the time it takes to see the cytopathic effect and virion release in *in vitro* experiments. HCMV however seems to have a moderately quicker replication rate *in vivo* with a doubling time of approximately one day, indicating that it may not be the replicative cycle itself which is slow, but possibly other factors also (Emery et al., 1999, Emery, 2001).

1.3 HCMV Genome

The class E structure (**Figure 1.1**) of HCMV's genome is organized into 2 unique regions-unique long (UL) and unique short (US) sequences that are each flanked by inverted repeats, referred to as TRL/IRL and IRS/TRS (terminal repeat long/short and internal repeat long/short). This results in an overall genome configuration of TRL–UL–IRL–IRS– US–TRS (Chee et al., 1990, Mocarski and Courcelle, 2001). The biological reason for this class E type structure remains unknown. HCMV's genes are named according to their position in the genome and are generally sequential, for example US1 to designate the 1st gene in the US region of the genome, however this can be disturbed due to the historical designation of genes prior to this agreement (Spaete et al., 1994, Chee et al., 1990).

HCMV has the largest genome of all the characterized herpesviruses and has evolved to produce a larger capsid than other herpesviruses, with a diameter of 130 nm (rather than the 125 nm of the HSV-1 capsid), however its DNA is still tightly packed within it (Butcher et al., 1998, Bhella et al., 2000). The genome size of HCMV varies with strain but ranges from 230 to 235 kDa (Mocarski et al., 2013). The genome of HCMV contains over 200 open-reading frames (ORFs) to date and encodes at least 170 protein-coding genes, many of these with unknown functions (Davison et al., 2003a, Dolan et al., 2004, Zhang et al., 2007, Gatherer et al., 2011). Core genes are found across all three subfamilies of human

herpesviruses (α, β and γ) and often encode genes that are essential for viral growth, such as proteins of the DNA polymerase complex and the major capsid protein (MCP)(Davison, 2007a). Only 45 of HCMV's genes are required for replication in fibroblasts, leaving the vast majority of genes having the potential to promote virulence and/or avoid the host's defences (Wilkinson et al., 2015, Dunn et al., 2003b, Yu et al., 2003).

Also present in the genome are cellular homologues that indicate that HCMV originally acquired many of its genes by 'gene capture', where recombination occurred from the host in a lateral transmission. Cellular homologues are common in herpesviruses and gene capture has resulted in the addition of functions in cell signalling, apoptosis and immune regulation (Davison, 2011). This has even occurred independently in various lineages during herpesvirus evolution, including with human interleukin-10 (*IL-10*) that occurred on 3 occasions across 2 sub-families in the genera *Cytomegalovirus*, *Lymphocryptovirus* and *Percavirus* (Davison, 2011). Capturing host anti-inflammatory cytokines could be advantageous to the virus as they can reduce the immune response to viral infections. HCMV has a broader selection of captured genes than other herpesvirus members, and also uses gene duplication as a way to produce gene families that can diverge in function (**Section 1.8**). HCMV is also unusual as it has a number of hypervariable genes that exist as different variants and can be highly divergent, thought to have been caused by ancestral immune selection in different human populations (Davison, 2011). Due to the possibility of co-infection with multiple strains, recombination is also evident between current HCMV strains which will similarly affect their evolution. As well as regions encoding proteins, HCMV also encodes three non-coding RNAs and 14 miRNAs (Davison, 2011). Multiple anti-sense RNAs are also known to be transcribed but their function is not yet clear (Gatherer et al., 2011).

1.3.1 HCMV strains

The main HCMV strains used in laboratories across the world are AD169, Merlin, TB40/E, Toledo, TR and Towne (**Table 1.3**). The phylogenetic relationships of some of these strains is shown in **Figure 1.3**. Different strains of HCMV can have varying genome length and vary in the copy number of terminal and internal repeats sequences, as well as containing mutations and deletions (**Table 1.3**). AD169 has been widely used in research since its original isolation in ~1955 from the adenoids of an infected child and was the first strain of HCMV to be fully sequenced (Chee et al., 1990, Rowe et al., 1953). Towne has also been

extensively used since its isolation back in 1970 (Plotkin et al., 1975). Both these strains are termed 'laboratory strains' or 'high passage strains' as they have been passaged extensively *in vitro* which has caused them to acquire mutations and they may no longer represent the clinical strain they originated from (Wilkinson et al., 2015, Cha et al., 1996). Numerous point mutations and deletions that have accumulated have changed not just the genetic content but also the biological properties of the viruses, resulting in AD169 and Towne becoming easier and quicker to cultivate than clinical isolates. It has however also resulted in them being unable to replicate in several cell types, including epithelial and endothelial cells (Ryckman et al., 2006) and they have been altered so much that they now cause little or no virulence when introduced to seronegative individuals (Just et al., 1975; Plotkin et al., 1976; Neff et al., 1979 as cited by Sijmons et al. (2014)).

Other isolates are termed 'low-passage strains', and these are much closer to HCMV isolates in the population but are not readily propagated in cell culture. This includes Toledo which has been passaged much less than the AD169 and Towne and still retains its virulence, producing the normal primary symptoms of a mononucleosis-type illness in the exposed individuals (Quinnan et al., 1984 as cited by Sijmons et al. (2014)). Merlin is a low passage strain originally isolated from a congenitally infected child (Davison et al., 2003a) and is often referred to as a 'clinical strain' as the cloned viral genome was restored to match the sequence found in the original clinical sample, prior to being passaged in cultured cells (Stanton et al., 2010). Merlin is the first strain to accurately represent a fully 'wildtype' clinical strain, and was achieved using a bacterial artificial chromosome (BAC) in which mutations gained *in vitro* were repaired. The BAC is an invaluable tool that allows for a reliable genetically stable source of 'wildtype' or 'clinical'-type virus and is used to avoid over-passaging the virus and to prevent mutations that would alter its genome (Stanton et al., 2010). The only difference to the clinical form of this strain is two genes (*RL13* and *UL128*) that are interrupted by point mutations that were rapidly selected in cell culture (Davison et al., 2003a, Dolan et al., 2004, Stanton et al., 2010, Dargan et al., 2010). This is not uncommon, with both *RL13*, and the UL128 locus (UL128L, consisting of genes *UL128*, *UL130* and *UL131A*) consistently and rapidly affected by mutations (Murrell et al., 2013). Generally the first mutation to be selected for in cell culture impacts *RL13*, with the UL128L mutating shortly afterwards. Mutations in the unique long/*b'* (UL/*b'*)

Table 1.3: A summary of the main research strains of HCMV, how they were originally isolated and the main mutations or alterations in their genomes.

Strain ◊	Type	Main differences†	Isolated from	References
Merlin	Low passage strain	The BAC derived virus has a fully wildtype genome, except for point mutations in *RL13* and *UL128*. Almost 100% identity to original clinical sample	Isolated from fibroblasts by 3 passages from a congenitally infected child's urine sample (Cardiff Diagnostic Virology laboratory)	(Wilkinson et al., 2008, Dolan et al., 2004, Wilkinson et al., 2015)
AD169	High passage lab-adapted strain	Deletion of 15 kbp (~19 genes) from the UL/*b*' region (UL133-UL151), acquired duplications of *RL1-14*, and mutations in *RL13* and *UL128L*. Altered tropism and virulence.	Isolated from the adenoids of a 7-year-old, and passaged 14 times in human fibroblast cells (National Institutes of Health, Bethesda, MD, USA)	(Rowe et al., 1953, Bradley et al., 2009, Cha et al., 1996, Sijmons et al., 2014).
Toledo*	Low passage strain	Some UL/*b*' region genes are inverted. BAC strain is missing *US2-US11*. Not representative of wild-type. Some variants don't replicate in endothelial cells.	Isolated from the urine of a congenitally infected child (Stanford University). BAC clone was a plaque-purified derivative	(Cha et al., 1996, Quinnan et al., 1984, Baldanti et al., 2003, Wilkinson et al., 2015)
Towne*	High passage lab-adapted strain	Deletion of ~13 kbp from UL/*b*' region including the region *UL144-UL151*, plus duplication of sequences within *UL1* as well as *a* and *b* repeats.	Isolated in 1970 from the urine of a 2-month old congenitally infected infant, and passaged in human fibroblasts	(Cha et al., 1996, Murphy and Shenk, 2008, Plotkin et al., 1975).

Strain	Type	Main differences	Isolated from	References
TR	Low passage strain	BAC clone is missing *US2-US5*.	Isolated from an AIDS patient with retinitis ("patient 2").	(Smith et al., 1998, Murphy et al., 2003)
TB40/E	Low passage strain	Mutations in UL/*b*' region including a frameshift in *UL141*. One derivative additionally lacks *UL145* and *UL144*. Highly endotheliotropic strain with release of high titres	Originally derived from a mixed population of mutant variants from a throat wash of a bone marrow transplant recipient by propagation in fibroblasts	(Sinzger et al., 2008, Wilkinson et al., 2015, Dolan et al., 2004, Tomasec et al., 2005)

◊ There are however multiple variants of most HCMV strains which can have varying properties and mutations

† Non-exhaustive list

* In Towne and Toledo BAC strains vector cassette sequences replaced the US1-US11 and US2-US11 genes respectively
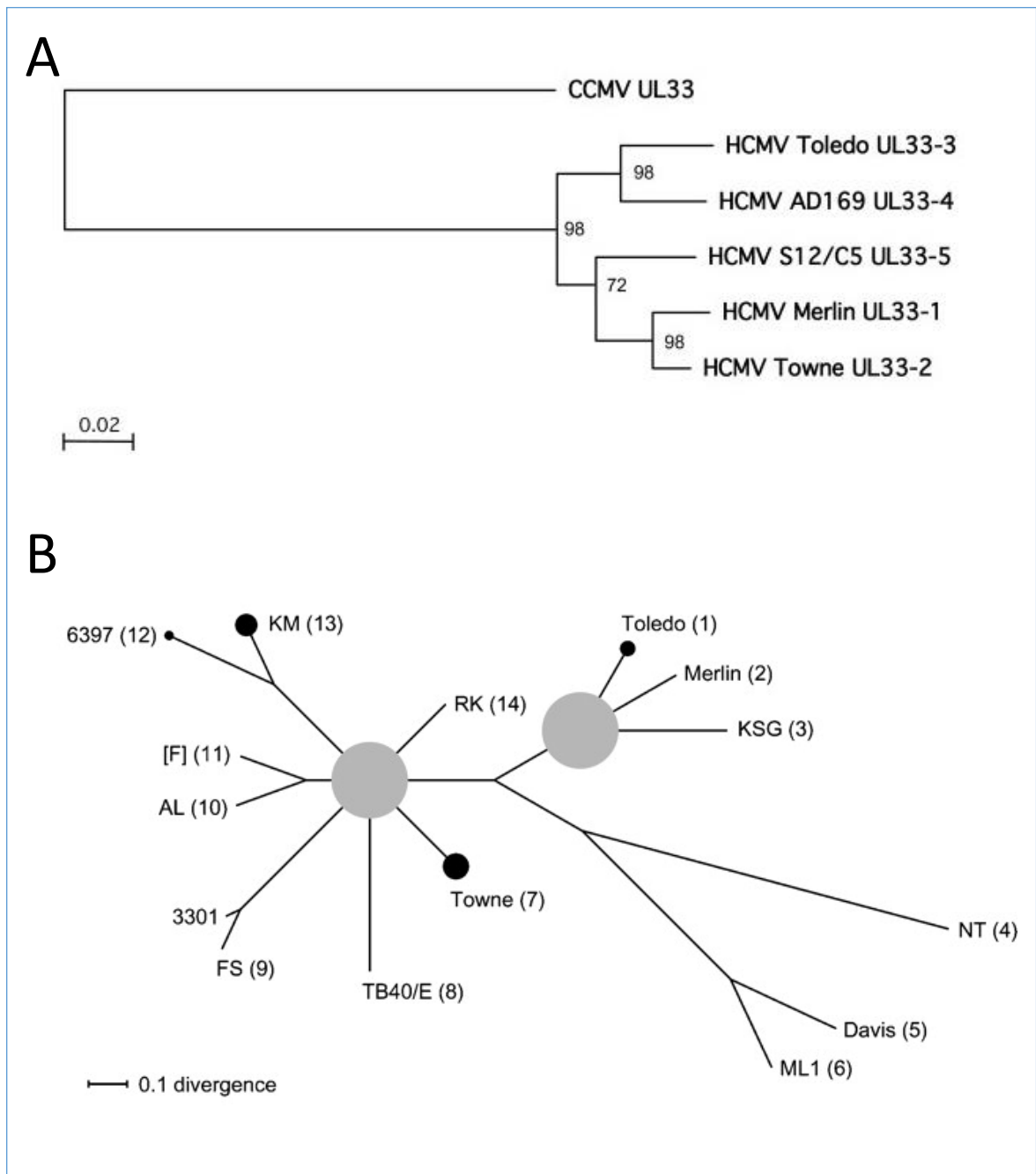
Figure 1.3: Phylogenetic analysis of several HCMV strains. A) Phylogenetic analysis by Deckers et al. (2009) of the UL33 proteins of multiple HCMV strains, using chimpanzee cytomegalovirus (CCMV) as outlier. The branch length is proportional to evolutionary distance. **B)** Phylogenetic analysis by Dolan et al. (2004) of the UL146 proteins of different HCMV strains. Grey circles indicate regions of unresolved branching order, and black circles indicate sets of closely grouped strains. The scale bar indicates divergence as substitutions per amino acid site.

region also occur but take longer and generally take place between the *UL140* and the *UL145* region (Murrell et al., 2013, Stanton et al., 2010, Dargan et al., 2010). Other genome alterations during cell culture are more strain specific and include a variety of mutations, deletions, duplications, differences in repeat regions and re-arrangement (Murphy and Shenk, 2008, Cha et al., 1996). AD169 and Towne for example both have extensive deletions in the UL/*b'* region due to their extensive passaging. The Towne strain is missing a ~13 kbp DNA segment from this region, including genes *UL144-UL151*, and the AD169 strain is missing a 15 kbp segment (~19 genes) including *UL133-UL151* (Cha et al., 1996) (**Table 1.3**).

Aside from the commonly affected UL/*b'* region, analysis of AD169 variants from 3 laboratories revealed that numerous other genetic changes had accumulated over time during passage of this virus. This also caused differences between the variants, and the situation is similar with the Towne strain (Wilkinson et al., 2015, Bradley et al., 2009). Both AD169 and Towne strains have acquired a multigene repeated sequence, with AD169 containing duplications of *RL1-14* (termed *TRL1-14* and *IRL1-14* to differentiate) and Towne underwent duplication of sequences within *UL1* as well as *a* and *b* repeats (Murphy and Shenk, 2008). A portion of the UL region of AD169, are present in the Toledo strain, but in an inverted orientation, with a shortened *b'* repeat sequence (Cha et al., 1996). Sporadic mutations can also occur across other regions over time and a few of the genes found to have been previously affected include *RL5A, UL36, UL131A, RL1, UL42, UL43* (Wilkinson et al., 2015, Sijmons et al., 2015).

Alongside Merlin there is now also a BAC-derived low passage TR strain, originally isolated from an AIDS patient with retinitis (Smith et al., 1998), as well as a TB40-E variant (TB40/E) strain, originally derived from a bone marrow transplant recipient as a mixed population of mutant variants (Wilkinson et al., 2015, Sinzger et al., 2008). The virus derived from BAC TB40/E can be grown to exceptionally high titres and is able to infect a wide range of cell types for a culture strain (Sinzger et al., 2008), however it still contains mutations in some key genes (Wilkinson et al., 2015, Dolan et al., 2004). Most BAC derived constructs also contain an incorporated cassette that replaces genes in the US2 and US6 gene families, with Towne and Toledo BAC-derived strains missing *US1-US11* and *US2-US11* genes respectively (Wilkinson et al., 2015)(**Table 1.3**).

1.4 HCMV Life Cycle

The HCMV life cycle *in vivo* is classically divided into productive lytic infection **(Section 1.4.1)** and non-productive infection termed 'latency' **(Section 1.4.2)**, although the reality may be much more complex. As herpesvirus infections are lifelong, initial exposure to the virus establishes systemic acute infection, but the virus must also establish sites of long-term persistence. All stages of the HCMV lifecycle can greatly alter the proteome of the cell surface, with changes in abundance occurring in 24% of the host proteins at 72 hours post infection (hpi) (Jean Beltran and Cristea, 2014).

1.4.1 HCMV Productive Lytic Infection

During the productive HCMV infection life cycle **(Figure 1.4)**, virions enter the host cell and initiate a multi-step process that involves delivery of virion components to the nucleus, initiation of transcription followed by a cascade through early to late phase gene expression, ultimately leading to the production of new virions and their egress from the cell. HCMV can replicate in a vast array of cell types *in vivo* and *in vitro*, and this broad tropism is important for viral spread within and between hosts.

1.4.1.1 Tropism

HCMV can target a broad range of cell types during infection including fibroblasts, epithelial cells, endothelial cells (ECs) and smooth muscle cells (Sinzger et al., 1995). Epithelial cells and ECs are found across a multitude of organs and tissues, with bone marrow, lung and gastrointestinal ECs frequently infected (Sinzger et al., 1995, Myerson et al., 1984). They can facilitate viral spread to and from the environment and other hosts, as well as facilitating the spread of HCMV to other organs within the same host (Söderberg-Nauclér and Nelson, 1999, Waldman et al., 1995, Hahn et al., 2004). Vascular ECs for example may allow HCMV to spread through the bloodstream, aiding spread and dissemination (Waldman et al., 1995, Grefte et al., 1993). EC subtypes can have differing gene expression and differ in their susceptibilities to HCMV (reviewed in Jarvis and Nelson (2007)), with some EC subtypes even being suggested as sources of latency (Fish et al., 1998). Other commonly infected cells are macrophages and fibroblasts, with HCMV occasionally found in neuronal cells, smooth muscle cells and hepatocytes (Söderberg-Nauclér and Nelson, 1999). For *in vitro* HCMV growth and experimentation, fibroblasts, endothelial cells and epithelial cells are most commonly used.
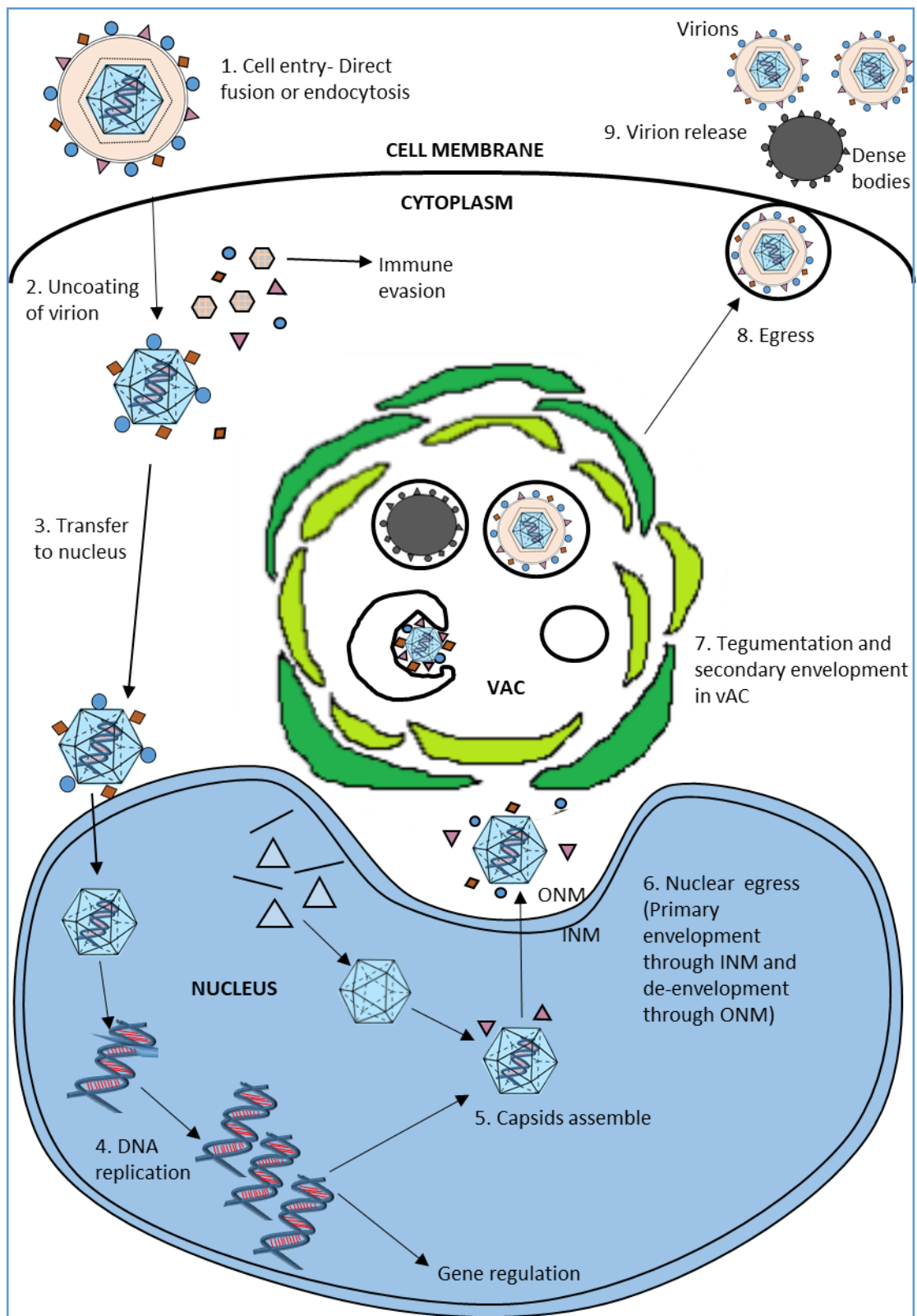
Figure 1.4: Simplified diagram of HCMV's lytic life cycle. Briefly, HCMV binds to cellular receptors and enters the cell by endocytosis or direct fusion. The capsid and a few tightly-bound tegument proteins are transported to the nucleus where viral DNA replication occurs. Capsids are assembled and the viral genome inserted, and the initial tegument is added. The capsids then undergo nuclear egress, with primary envelopment through inner nuclear membrane (INM) and de-envelopment through outer nuclear membrane (ONM). In the virion assembly compartment (vAC), final tegument is added and secondary envelopment occurs, before the virions are transported out of the cell via the exocytosis pathway.

HCMV transcription differs between cell types, with the transcriptome of the Merlin strain shown to be different in fibroblasts, epithelial cells and astrocytoma cells over time (Towler et al., 2012). Genes that are down-regulated include those that are involved in genome replication, virion assembly, and virion maturation and release stages (Towler et al., 2012) which may account for the differing growth kinetics in the different cell types.

1.4.1.2 HCMV Entry

The first stage of the life cycle process is that the virus must enter the host cell by membrane fusion, mediated by the binding of HCMV's envelope glycoproteins to host cell receptor(s). Glycoprotein B (gB) mediates HCMV entry by acting as a fusion protein, which is required for virus entry and cell-to-cell spread, although not absolutely required for virus attachment (Isaacson and Compton, 2009, Wille et al., 2013). There is some evidence for the direct binding of gB to the platelet-derived growth factor receptor (PDGFR) and the epidermal growth factor receptor (EGFR), although PDGFR does not appear to be required for standard HCMV entry and EGFR is not required for cellular expression of HCMV-essential genes (Soroceanu et al., 2008, Vanarsdall et al., 2012, Wang et al., 2003). There is also the possibility that gB binds integrins and activates integrin-specific signal transduction pathways (Feire et al., 2004).

Glycoproteins gH and gL are also needed for both cell entry and for cell-cell fusion; and alongside gB they form the minimal HCMV fusion machinery, where they possibly trigger gB for fusion (Wille et al., 2013, Vanarsdall and Johnson, 2012, Isaacson and Compton, 2009). These entry proteins can be seen as tropism determinants, and gH and gL form the core components of 2 distinct complexes that are required for viral entry in different cell types. The 'pentameric complex' consists of gH, gL, pUL128, pUL130A and pUL131 (gH/gL/UL128-131) and it is required for viral infection of endothelial, epithelial and myeloid cell types, whereas the 'trimer complex' consisting of gH, gL and gO (gH/gL/gO) is required for entry into fibroblasts (Wang and Shenk, 2005b, Wang and Shenk, 2005a, Hahn et al., 2004). Recent updates suggest however that it may in fact only be gH/gL that is required in the virion for fibroblast entry in the TR strain at least, with gO just required as a chaperone (Wille et al., 2010). Other complexes or additional proteins may also be involved (Caló et al., 2016). A viral regulator of these complexes has also been found to affect cell tropism, with UL148 suggested to favour the gH/gL/gO complex by competing with pUL128 for the partially assembled gH/gL complexes, regulating the relative amounts of the two gH/gL

complexes (Li et al., 2015). Some strains have selected for mutations in these entry complex genes due to passaging in certain cell types and therefore can no longer propagate in other cell types. Mutations in *UL128*, *UL130* or *UL131* are capable of abolishing endothelial cell tropism (Hahn et al., 2004, Akter et al., 2003), and repairing the mutated *UL131* gene in AD169 restores its ability to infect both epithelial and endothelial cells (Wang and Shenk, 2005a).

Glycoproteins gM and gN are found in high abundance in the virion envelope, are both essential for HCMV replication, and form a complex (gM/gN) (Mach et al., 2000, Krzyzaniak et al., 2007, Varnum et al., 2004). It is not clear whether the gM/gN complex promotes virus entry, but it has been suggested that they can promote virus assembly and envelopment of capsids (Krzyzaniak et al., 2007, Mach et al., 2007).

Once the glycoproteins have bound, entry of the virus occurs either by fusion or endocytosis, depending on the cell type and the conditions. In epithelial and endothelial cells, endocytosis occurs and as the virus enters it becomes enclosed in the plasma membrane in an endosome. Escape from the endosome requires acidification, with the low pH triggering the viral membrane proteins to fuse with the endosomal membrane (Wang and Shenk, 2005b, Bodaghi et al., 1999). In fibroblasts however, entry is by direct fusion with the plasma membrane and virions do not become enclosed in endosomes, so direct entry is pH-independent (Compton et al., 1992).

Tegument proteins from the input virus can regulate some of the cellular pathways which are predicted to be involved in the final steps of entry, and the stages of transporting the uncoated capsid to the nucleus (reviewed by Kalejta (2008)). The interaction of gB (pUL55) and gH (pUL75) with host cell surface receptors upon viral entry also activates cellular transcription factors, including nuclear factor kappa B (NF-κB) and specificity protein 1 (Sp1), as well the interferon (IFN) response by strongly inducing many IFN-stimulated genes (Huang and Johnson, 2000, Yurochko et al., 1997, Boehme et al., 2004, Simmen et al., 2001). HCMV encodes proteins to downregulate these antiviral immune responses, including the input protein pp65 (ppUL83) which can dampen and the induction of pro-inflammatory cytokines and IFN-responsive genes (Browne and Shenk, 2003). The NF-κB pathway however is still required for the transactivation of the major immediate early promoter (MIEP) so its modulation must be balanced (DeMeritt et al., 2004, Browne and Shenk, 2003). NF-κB activation can start as early as 15 mins after HCMV entry, with the pathway subsequently activated within ~24 hpi (Yurochko et al., 1997).

1.4.1.3 Nuclear Transport

After membrane fusion, some tegument proteins are released while pUL47 and pUL48 remain tightly bound to the capsid where they are thought to interact with the host dynein-microtubule system in order to transport the viral capsids to the nuclear pore (Liu and Zhou, 2007, Ogawa-Goto et al., 2003). Viral DNA enters the nucleus through the nuclear pore complex (NPC) where pUL47, pUL48, pUL69 and the major capsid protein (MCP; pUL86) together facilitate the release of the viral DNA from the capsid into the nucleus (Bechtel and Shenk, 2002, Kobiler et al., 2012). Some tegument proteins inhibit the initial immune response or regulate viral gene expression, such as pp65 (reviewed by Kalejta (2008)). Others can migrate to the nucleus independently of the nucleocapsid, including pp71 (ppUL82) which subsequently targets the human death-domain associated protein hDaxx (Woodhall et al., 2006). Other tegument proteins traffic to different sub-cellular locations, such as pp28 which localises to the vAC (**Section 1.4.1.6**) (Sanchez et al., 2000b).

Once inside the nucleus, the genome requires interactions with cellular histones which can be modified by various enzymes that can remodel the chromatin, such as histone deacetylase 1 (HDAC1). This enzyme is a member of the nucleosome remodelling and deacetylase (NuRD) protein complex, and it can regulate immediate-early (IE) gene expression (Terhune et al., 2010, Groves et al., 2009, Reeves et al., 2006). Once in the nucleus, viral genome transcription and replication can be initiated.

1.4.1.4 Viral Gene Expression

A large number of viral genes are expressed across lytic replication, which are not all expressed simultaneously, but in a cascade of expression. Traditionally there has been 3 main cascades, immediate early (IE/α), early (E/β) and late (L/γ) cascades.

Immediate early gene expression

Immediate early (IE, α) genes are the first to be transcribed and these are triggered after cell entry, usually appearing within 1 hour post infection (hpi) and peaking around 4-8 hpi. Initiation of IE gene expression depends upon pre-existing machinery so is not reliant on de novo viral protein synthesis and is also sensitive to the cell cycle phase (Stinski et al., 1983, Salvant et al., 1998, Stenberg et al., 1984). IE transcripts therefore accumulate in the presence of cycloheximide or other protein synthesis inhibitors (Chambers et al., 1999). IE gene expression is mapped to 4 regions of the genome, with the majority of expression occurring at the IE1/IE2 locus, giving rise to the major IE1 and IE2 products, with their

expression controlled by the major immediate early promoter (MIEP) (Wilkinson et al., 1984, Akrigg et al., 1985, Stenberg et al., 1984). These 2 gene products share an 85 amino acid domain which splices to UL123 or UL122 respectively (Stenberg et al., 1984, Stinski et al., 1983, Akrigg et al., 1985). IE genes can function as transcription factors and by definition can instigate the onward transcriptional cascade (White and Spector, 2007, Guetta et al., 2001, Stinski et al., 1983). IE2 is vital for activating the subsequent stages and is essential for viral replication and has many functions (Sarisky and Hayward, 1996, Iskenderian et al., 1996). IE2 is also implicated in auto-regulation of the MIEP, with IE1 functioning as an accessory protein, having an indirect enhancement on transcription (Wilkinson et al., 1998, Reeves et al., 2006). Aside from IE1 and IE2, the other 3 regions of IE/α genes are UL36-UL38, TRS1-IRS1 and US3 (Colberg-Poley, 1996).

Activating the viral cascade is important, but detailed studies reveal that IE genes are primarily required to counter intrinsic, innate and adaptive immune defences. pIE1 for example disrupts PML-bodies; pIE1, pIRS1 and pTRS1 counter the interferon response, pUL36 and pUL37 have anti-apoptotic functions, and pUS3 sequesters MHC-1 (Colberg-Poley, 1996, Noriega et al., 2012b, Skaletskaya et al., 2001, Goldmacher, 2005, Child et al., 2004, Wilkinson et al., 1998). Although not an IE protein itself, pp71 (UL82) is a virion protein that acts as a transactivator and is required for IE gene expression, and functions through binding with and degrading hDaxx to relieve hDaxx and HDAC-mediated silencing of MIEP (Cantrell and Bresnahan, 2006, Woodhall et al., 2006).

Early gene expression

Early (E/β) genes are next to be transcribed, with their gene expression triggered by functional IE genes, especially IE2, and are therefore triggered ~8-12 hpi. Early genes are transcribed even in the presence of a viral DNA synthesis inhibitor and are required for initiating viral genome replication (Chambers et al., 1999). Genes in this cascade include genes that play a role in viral DNA synthesis and processing, as well as immune modulators, anti-apoptosis genes and cell cycle regulators (Huang and Johnson, 2000, White and Spector, 2007). Most early genes accumulate gradually over the course of infection, and generally continue through the late phase. Classically these genes were split into 2 subsets, β1 (early) and β2 (early-late) genes as defined by a difference in their expression patterns, with β2 genes commonly partially inhibited by viral DNA synthesis inhibitors such as Phosphonoformate (PFA) (Chambers et al., 1999).

Late gene expression

Late (L/γ) genes are the last to be produced and their expression is highest after viral DNA replication has begun (~24 hpi), their expression is dependent upon the expression of E genes, and they are partially or completely inhibited by the use of viral DNA synthesis inhibitors (Chambers et al., 1999). This cascade subset includes mostly structural virion proteins, capsid maturation proteins or proteins that play roles in virion maturation and egress (Brinkworth and Thorn, 2013, Mocarski et al., 2013). Classically, late (γ) genes were divided into the 2 sub-categories of leaky-late (γ1) and true-late (γ2) proteins depending on their pattern of expression and their dependence on viral DNA synthesis, with true-late genes expressed exclusively after DNA replication. Their pattern and timing can differ with cell type however.

Temporal kinetics classification

Another gene classification system has also been created by studying the temporal protein analysis of 139 canonical HCMV proteins and 14 non-canonical ORFs, clustering them based on their time of expression and expression patterns (Weekes et al., 2014). This demonstrated that all detected HCMV proteins clustered into 5 distinct cascades, classified as temporal classes Tp1-5 the average expression patterns of which are depicted in **Figure 1.5**. This temporal class classification system is believed to be more definitive and accurate than the previous system which had studied proteins in a range of different viral strains and cell types. It does however have a good correlation with the classical classes with the known late proteins UL99, UL94, UL75, UL115, and UL32 all correspondingly classified as Tp5 for example (Weekes et al., 2014). The use of PFA also demonstrated a good correlation between the 2 systems, generally having little effect on Tp1 or Tp2 proteins, partially inhibiting Tp3 and Tp4 proteins, and completely or almost completely inhibiting the majority of Tp5-class protein s (Weekes et al., 2014). This temporal classification system also has a good correlation with other classification systems, with ten of the thirteen Tp1 proteins being categorised into the equivalent temporal class (Tr1) by their mRNA (Weekes et al., 2014, Stern-Ginossar et al., 2012). Proteomics did however reveal that eight proteins (UL27, UL29, UL135, UL138, US2, US11, US23 and US24) were expressed earlier in infection than previously understood (Weekes et al., 2014).

Tp1 proteins are expressed with a similar pattern to IE proteins, demonstrating relatively high levels of expression from ~6 hpi (**Figure 1.5**). Both Tp2 and Tp3 proteins show a similar pattern to early and early-late proteins (Weekes et al., 2014), with Tp2 proteins tending to accumulate rapidly through early time points, staying high and constant throughout
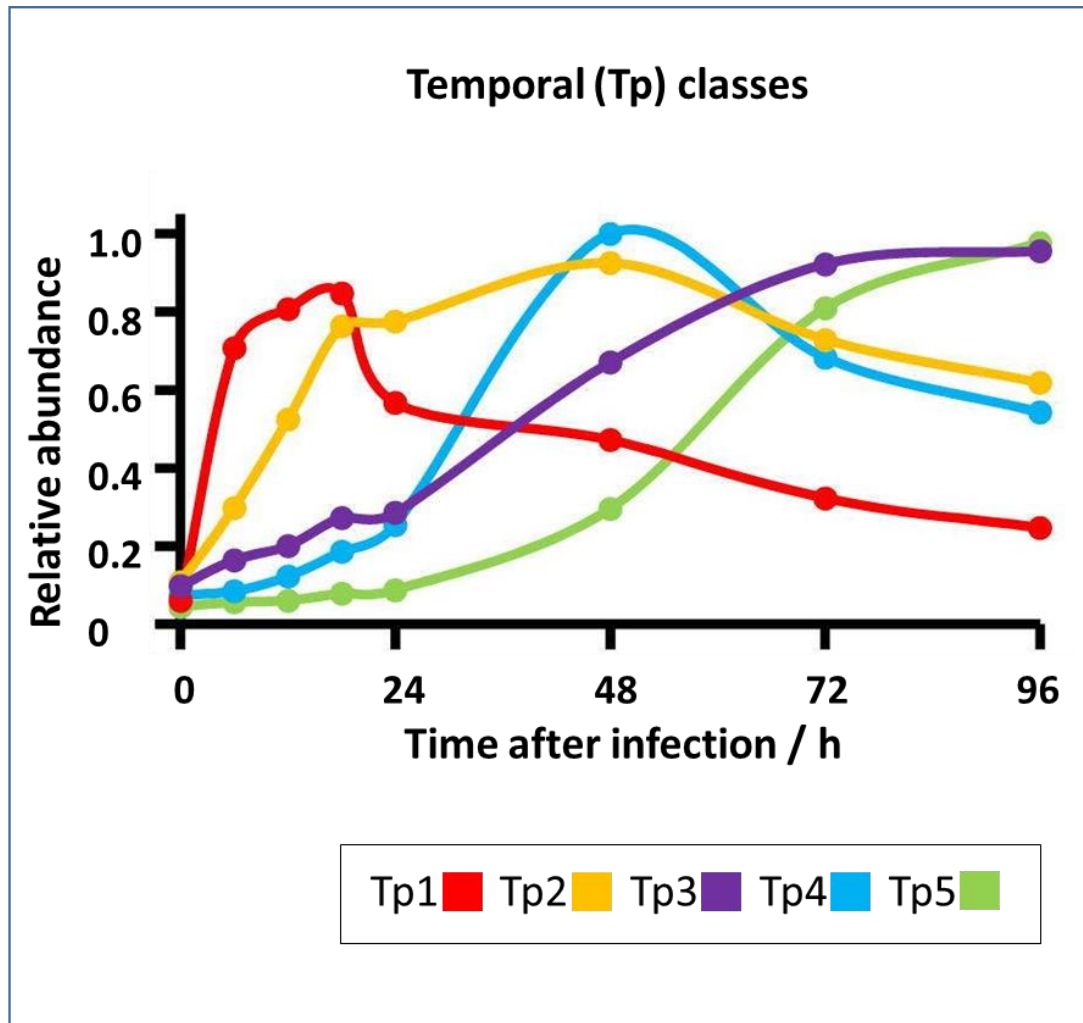
Figure 1.5: The 5 Temporal Classes of HCMV Gene Expression. HCMV gene expression profiles were classified into 5 distinct classes depending on their gene expression patterns and timings. Figure from Weekes et al. (2014). They used k-means clustering to assign all quantified HCMV proteins to temporal classes. Shown are the average temporal profiles of each class**.**

infection. An example of a Tp2 protein is the DNA polymerase pUL54 which is produced early and peaks ~24 hpi, levelling off over the infection time course (Weekes et al., 2014). Tp3 proteins tend to accumulate slower but steadily, peaking at late time points, such as with pUL71 (Weekes et al., 2014). Tp4 proteins have a distinct profile, with maximum expression at ~48 hpi and low/lower expression at other time points (Weekes et al., 2014). These proteins seem to cluster in the genome and includes the gpRL11 and gpRL12 proteins (Weekes et al., 2014). Tp5 proteins have minimal protein expression at 24 and 48 hpi and peak between 72 and 96 hpi, with PFA inhibiting 87% of Tp5 proteins (Weekes et al., 2014). pp28 (ppUL99) is a Tp5 protein (also categorised as γ2) and has low expression levels at 48 hpi, and peak expression at 96 hpi. All of HCMV's capsid proteins are also Tp5 proteins, including the major capsid protein (MCP, pUL86) (Weekes et al., 2014). The majority of late proteins were originally classified as γ1 proteins, but due to the redefining of the late protein subsets, most late proteins (85%) were re-classified as Tp5 proteins (Weekes et al., 2014), with this new classifications having a high level of correlation with the mRNA category defined as Tr5 (Stern-Ginossar et al., 2012).

1.4.1.5 Viral DNA Replication

Viral DNA (vDNA) synthesis does not occur until around 24 hpi in cell culture and is initiated by β2 (Tp3) gene expression. Once the viral DNA is released from the capsid into the nucleus, it circularises and likely forms into a chain of multiple copies of the DNA sequence linked together in a 'concatemer' (Pari, 2008). Replication takes place from the origin of lytic replication, OriLyt, which is the only functional replicator in the genome of HCMV. OriLyt is found in the UL region of the genome between *UL57* and *UL69* and is ~1500bp long and structurally complex (Anders et al., 1992, Pari, 2008). The majority of the proteins required for synthesis and processing of the viral DNA are expressed with early kinetics (Pari and Anders, 1993). pUL44 is responsible for bringing together the other components of the replisome complex which includes 11 loci in total, including genes encoding a DNA polymerase and DNA binding proteins. These loci include *IE1/2*, *UL36-UL38*, *UL54*, *UL57*, *UL70*, *UL84*, *UL101-2*, *UL105* and *IRS1/TRS1* (Pari and Anders, 1993). UL112-113 may also play a role in the recruitment of the core replication machinery proteins (Ahn et al., 1999). The pUL84 gene product is essential for viral DNA synthesis and productive infection and is suggested to act in concert with IE2-p86 to trigger viral DNA synthesis (Anders et al., 2007, Sarisky and Hayward, 1996).

Tens of thousands of viral genome copies can be produced per cell by 72-96 hpi during lytic replication in fibroblasts, however under 1000 genome copies per cell are produced in

epithelial (RPE-1s) and astrocytoma (U373MGs) cells, with total viral yields 1,000- and 10,000-fold lower respectively (Towler et al., 2012).

1.4.1.6 Virus Assembly and Structure

HCMV virions share the common herpesvirus structure of an envelope, tegument, capsid and a core (**Figure 1.6**). Assembly of the virus involves forming the capsid in the nucleus and inserting the DNA into the core of the capsid. The tegument is then formed around the capsid scaffold, and the particle is enveloped before egressing from the host cell.

Capsid assembly

HCMV capsids have the common herpesviruses shell structure of 100nm that consists of 162 capsomeres. Of these, 150 are hexameric and 12 are pentametric, with both pentons and hexons having a cylindrical shape (Chen et al., 1999). These are connected in groups of three by asymmetric structures on the capsid floor which form into the icosahedral capsid that has 20 triangular faces (Liu and Zhou, 2007). The capsid is formed of 4 main proteins- the major capsid protein (MCP/pUL86), the small capsid protein (SCP/pUL48A), the minor capsid protein (mCP/pUL46) and mCP binding protein (mCP-BP/pUL85) (Colberg-Poley and Williamson, 2013). The penton consists of five copies of the MCP (pUL86) and the hexon has six copies of MCP along with six copies of the SCP (pUL48.5) (Chen et al., 1999, Gibson, 1996). Proteins also required for capsid development include the proteinase precursor (pNP1, pUL80A) and related assembly protein precursor (pAP, pUL80.5) which help to transport the MCP to the nucleus (Wood et al., 1997, Plafker and Gibson, 1998). mCP-BP on the other hand contains its own nuclear localisation signal so appears to transport itself and its partners mCP and SCP to the nucleus separately (Plafker and Gibson, 1998). Once the capsid is fully formed, the cleavage of pAP and pNP1 by the viral protease (Pr) disrupts their interactions with the MCP and allows the internal scaffolding to be removed (Welch et al., 1991, Yu et al., 2005). This makes space in the capsid for the viral genome, with pUL56 and pUL89 playing a role in the cleaving and packaging of the viral DNA through the portal protein (PORT/pUL104) formed on a single vertex of the capsid (Colberg-Poley and Williamson, 2013, Bogner et al., 1998). The timing of the protease is essential for DNA encapsidation and fully infectious virions (Yu et al., 2005) and to avoid incomplete capsids being formed. 'A' capsids are empty capsid shells, 'B' capsids contain scaffold but no DNA, and 'C' capsids are fully-formed and contain viral DNA and will subsequently become infectious virions (Tandon et al., 2015).

A

Viral dsDNA

Envelope glycoproteins

Capsid
(Icosahedron)

Lipid
envelope

Tegument

Tightly bound inner
tegument layer

B

Secondary
envelope

Vesicle for
egress

Capsid

Tegument

100nm
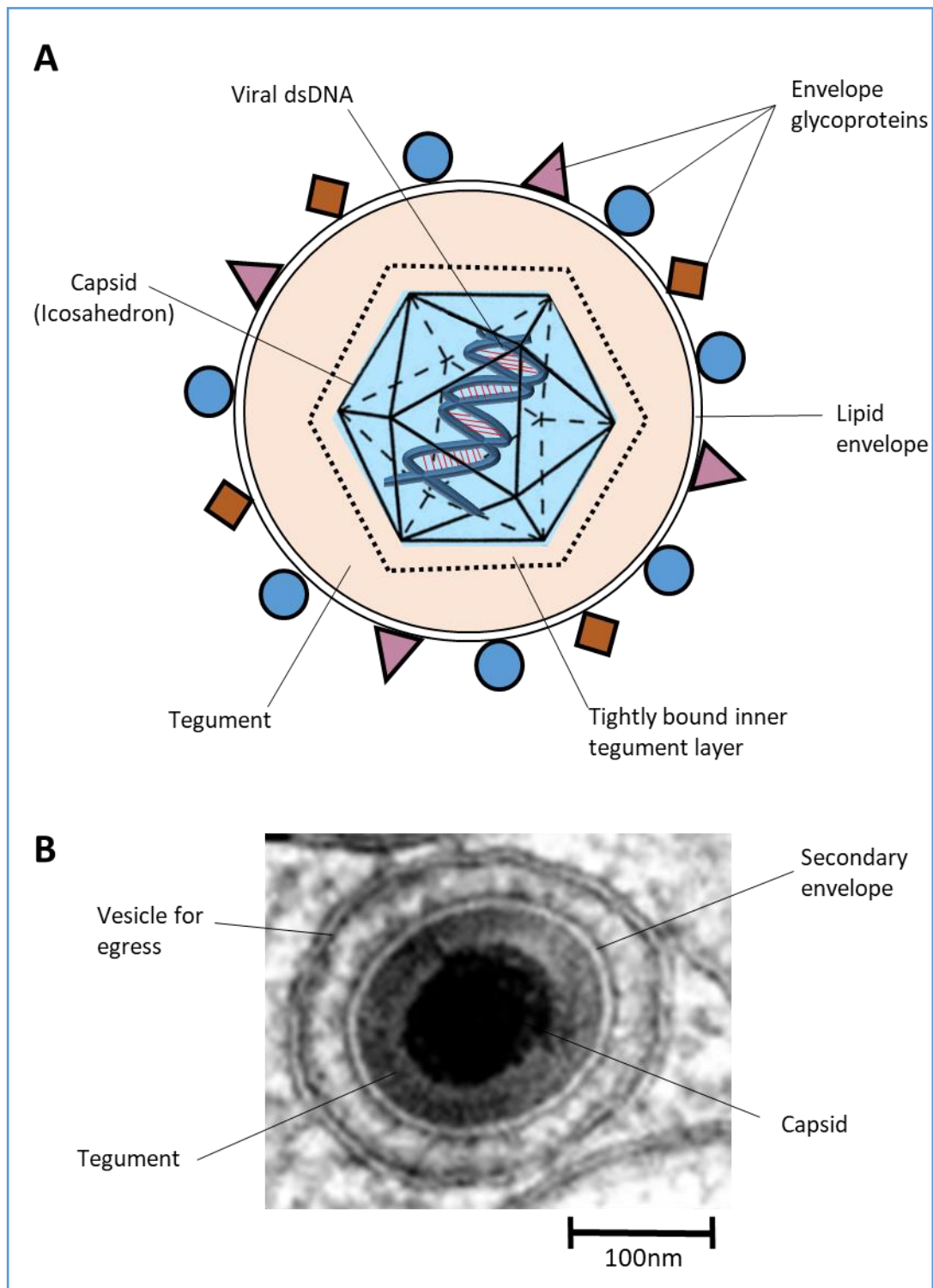
Figure 1.6: HCMV virion structure. **A)** Simplified version of the virion structure of HCMV including the 4 main components, the envelope, the tegument, the capsid and the core (area inside the capsid where the viral DNA is held). **B)** STEM tomography of the HCMV particle within a vesicle after secondary envelopment has taken place, taken from Schauflinger et al. (2013).

Addition of Tegument

Once the DNA is packaged, the addition of tegument can start. The tegument is a unique feature of herpesviruses and is found between the capsid and the envelope, with HCMV's tegument containing at least 38 viral proteins (Guo et al., 2010, Varnum et al., 2004). Some tegument proteins are conserved in all herpesviruses, with UL71, UL47, UL48, UL87, UL94, UL95, UL97, UL99, and UL103 all having homologs in HSV-1, EBV and KSVH (Guo et al., 2010). Despite this conservation, the tegument has no distinctive features (Guo et al., 2010, Varnum et al., 2004). The capsid acts as the scaffold to which the tegument is added (Chen et al., 1999). After previous debate as to the location of the tegument addition, it is now established that it first occurs in the nucleus and then continues in the virion assembly compartment (vAC) described below (Britt and Boppana, 2004). The localisation of protein expression can indicate at which site they are added to the maturing virion. pp65 (pUL83) for example translocates over time, migrating to the nucleus independently of the capsid and vDNA (Britt and Boppana, 2004) so it is likely that pp65 is added to the virion in the nucleus (Ahlqvist and Mocarski, 2011). The same is suggested for pp71 which is an important protein for the initiation of tegument assembly (Liu and Zhou, 2007, Ahlqvist and Mocarski, 2011). pp71 is suggested to provide structure by having a tight association with the nucleocapsid, along with proteins pp150 (ppUL32), pp28 (ppUL99) and pUL48 (Chen et al., 1999, Liu and Zhou, 2007). The location of where pp150 is added to the capsid is debated, but it is the most capsid-proximal tegument protein and is essential for tegument formation, and it is possible that it helps with the stability and transport of the capsid to the vAC (Hensel et al., 1995, Salsman et al., 2008, Meyer et al., 1997, Sanchez et al., 2000a, Homman-Loudiyi et al., 2003).

Additional tegument proteins are added in the vAC after nuclear egress has occurred. Many tegument proteins co-localise to the vAC during virion assembly, but it is unclear whether they are added to the virion at this location or whether they accumulate here due to their attachment to the capsid/initial tegument. Viral tegument proteins are generally phosphorylated and the tegument also contains cellular proteins as well as RNAs (Kalejta, 2008). These all accumulate during virion assembly and are likely to reside in the tegument. The tegument is also known to have DNA polymerase, protein kinase, and cellular topoisomerase for DNA replication (Huang and Johnson, 2000). By studying homologs in other herpesviruses, it has suggested that some of these outer tegument proteins bind to the membrane envelope (Guo et al., 2010).

Nuclear egress

With the initial tegument added in the nucleus, the nucleocapsids can egress through the nuclear membrane. This occurs in the part of the nuclear membrane that faces the vAC and a nuclear egress complex (NEC) is formed at the inner nuclear membrane (INM) (Alwine, 2012, Colberg-Poley and Williamson, 2013). It is composed of a type II membrane spanning component (NEC1, pUL50) and nuclear lamina-interacting component (NEC2, pUL53) which facilitate the capsids exit from the nucleus (Milbradt et al., 2007, Sharma et al., 2014). This occurs by disruption and re-modelling of the nuclear lamina by cellular and viral kinases and chaperones such as pUL97 (Marschall et al., 2005, Milbradt et al., 2007, Sharma et al., 2014). The NEC also appears to act as a quality control checkpoint, preferentially allowing DNA-containing capsids to egress over incorrectly formed ones (Tandon and Mocarski, 2012). It is suggested that the capsids are likely to be enveloped through the inner nuclear membrane (INM) and de-enveloped through the outer nuclear membrane (ONM), with capsids being released into the cytoplasm, and possibly transported via the microtubule-organizing center (MTOC) system to the vAC (Alwine, 2012, Das et al., 2007).

VAC Formation

During HCMV infection, the virus 'hijacks' the host cell's secretory machinery and arranges them in concentric cylinders known as the virion assembly compartment (vAC). This involves a multitude of host cell organelles, with Golgi and trans-Golgi apparatus circling the endosomal machinery, or at least vesicles that share properties with recycling endosomes (**Figure 1.7**) (Das et al., 2014, Das et al., 2007). There is also a network of microtubules (cytoskeletal filaments) radiating outward from the centre of the cVAC where the microtubule- organizing centre (MTOC) is located (Das and Pellett, 2011, Das and Pellett, 2007, Das et al., 2007, Sanchez et al., 2000a). This structure arises adjacent to the nucleus which causes the common 'kidney-shaped' nucleus frequently seen in HCMV infected cells. This is due to dynein pulling the nuclear membrane toward the MTOC, likely promoting a tight association with the vAC, with nuclear enlargement alleviating the nuclear breakdown that would occur otherwise (Alwine, 2012).

As well as host organelles, the vAC also contains tegument proteins including pp28 (ppUL99), pp65 (ppUL83), pp71 (ppUL82), and pp150 (ppUL32), and envelope glycoproteins such as gB (gpUL55), gH (gpUL75) and gL (gpUL115)(Sanchez et al., 2000b, Sanchez et al., 2000a, Homman-Loudiyi et al., 2003).

Key: Golgi; Trans-Golgi (TGN); Early endosomes; Microtubules
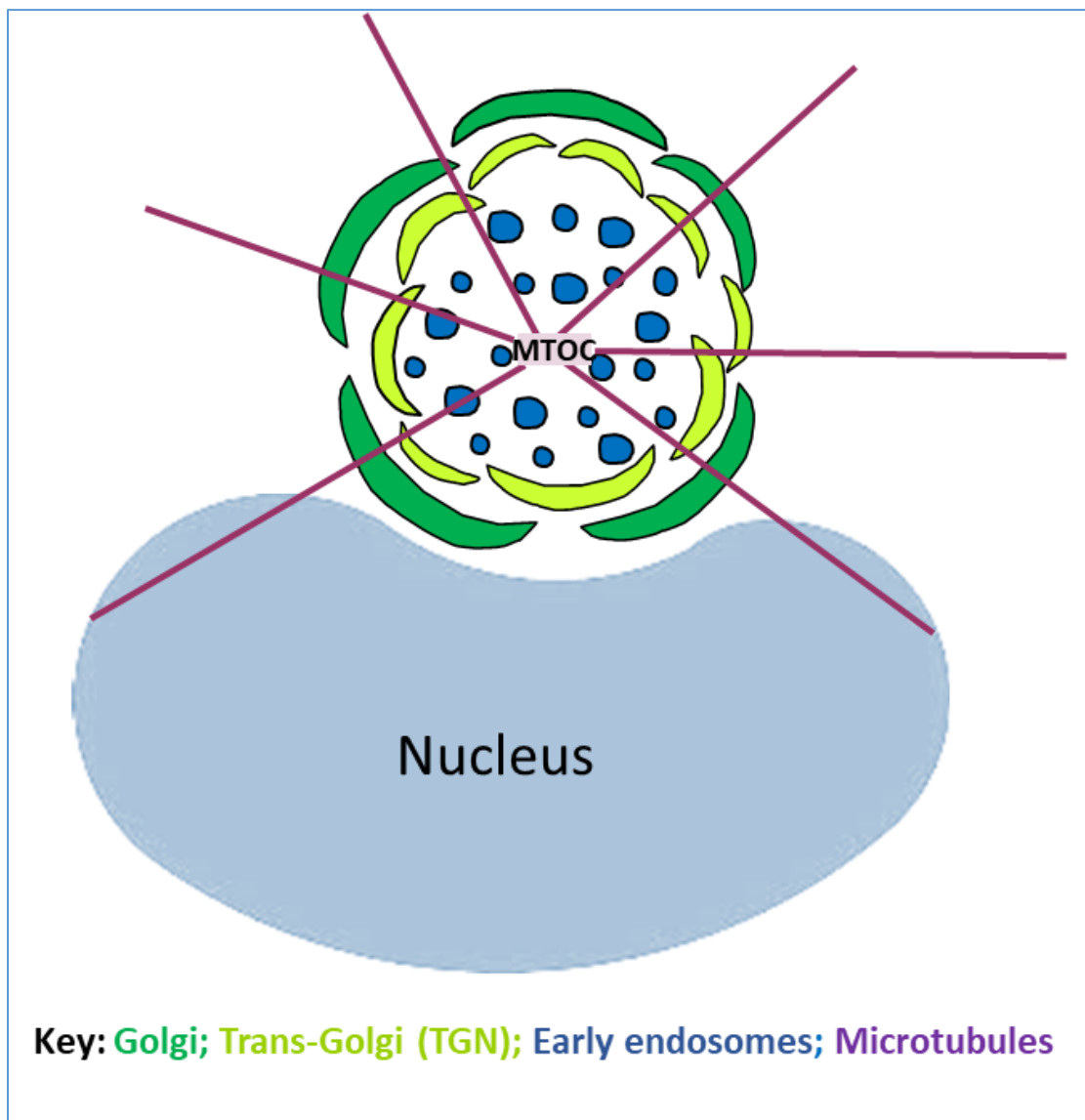
Figure 1.7: The virion assembly compartment (vAC). HCMV re-arranges the cellular secretory machinery and arranges them in concentric cylinders, including Golgi and trans-Golgi apparatus circling vesicles that share properties with endosomes. This forms around the microtubule-organizing centre (MTOC) from which a network of microtubules (cytoskeletal filaments) radiating outward from the centre of the cVAC. This structure arises adjacent to the nucleus and causes the common 'kidney-shaped' nucleus frequently seen in HCMV infected cells.

The vAC is a unique feature of β-herpesvirus-infected cells and is essential for HCMV virion assembly and egress (Das et al., 2007). This structure requires viral DNA synthesis to form, with pUL48, pUL94, and pUL103 vital for vAC development (Das et al., 2014). Little infectious virus is produced before the vAC is fully developed, which can take 3 to 4 days of re-modelling in fibroblasts by HCMV AD169, but can be quicker in other strains (Alwine, 2012, Das et al., 2007). Secondary envelopment is also believed to occur in the vAC, possibly in the early endosome-derived region (Das et al., 2014, Liu and Zhou, 2007, Hollinshead et al., 2012, Tooze et al., 1993). The reorganisation of host organelles also involves the rearrangement of the endoplasmic reticulum (ER), which becomes more condensed as cytoplasmic structures, compared to its typical diffuse staining pattern in uninfected cells (Buchkovich et al., 2009, Cavaletto et al., 2015).

1.4.1.7 Final Envelopment

The envelope is the outermost layer of the virus and contains most or all of the virion glycoproteins (Gibson, 1996). It is a triple layered structure and is acquired from altered cell membranes with the presence of lipids and numerous projections of glycoproteins (Homman-Loudiyi et al., 2003, Sanchez et al., 2000a). Previous studies suggested that these membranes could originate from the Trans-Golgi network (TGN), the ER-Golgi intermediate compartment (ERGIC), recycling endosomal compartments or vesicles derived from the above (Homman-Loudiyi et al., 2003, Cepeda et al., 2010). However, it is now generally accepted that the envelope proteins are first processed through Golgi/TGN, then exported to the plasma membrane and are retrieved by the endocytic system before being used to envelope the capsids (Hollinshead et al., 2012).

HCMV virions contain over 70 viral proteins as well as over 70 host cellular proteins (Varnum et al., 2004, Murrell, 2014), and the envelope of AD169 contains ~ 23 viral glycoproteins including gB, gH, gL, gM (gpUL100), gO (gpUL74), gp48 (gpUL4), gpTRL10, and gpUL33 (Varnum et al., 2004). Some proteins however are typically found in much higher concentrations than others, including gB which has around 800 molecules per AD169 virion (Varnum et al., 2004). Although there is a high degree of conservation between the virions and dense bodies of different HCMV strains, their compositions are flexible, and can differ across different strains and when grown in different cell types (Scrivano et al., 2011, Büscher et al., 2015, Murrell et al., 2013). The Merlin virion proteome shows many similarities to the virion of AD169, but the sensitivity of detection or strain differences revealed novel virion components of high confidence, some of which have homologs in RhCMV or MCMV virions (Murrell, 2014). The composition of the gH/gL entry complexes

can also be regulated, which can subsequently affect tropism as well as cell-cell transmission (**Section 1.4.1.1 and 1.4.1.2**) (Li et al., 2015, Luganini et al., 2017, Murrell et al., 2017, Wille et al., 2010). The membranes of dense bodies and enveloped virus appear to have the same composition, although dense bodies contain less cellular proteins (Büscher et al., 2015, Homman-Loudiyi et al., 2003).

1.4.1.8 Virion Egress

Mature enveloped infectious virions are released from the cells alongside non-infectious enveloped particles (NIEPs) and dense bodies (Tandon et al., 2015). Fully-formed virions contain C capsids with complete viral DNA and are therefore infectious. NIEPs on the other hand are enveloped particles which contain B capsids with either no or incomplete DNA which makes them non-infectious (Irmiere and Gibson, 1983). Dense bodies contain enveloped tegument proteins but lack capsids and viral DNA and are composed mainly of pp65 (Baldick and Shenk, 1996, Irmiere and Gibson, 1983, Büscher et al., 2015). Different amounts and percentages of each type of viral particles can be released by different strains and in different cell types, with AD169 overproducing NIEPs by ~10-fold in fibroblasts compared to other strains (Irmiere and Gibson, 1983).

After acquisition of the envelope, the particles exit the cell via the common cellular exocytosis pathway. This is facilitated by pUL103 (VEP), which controls the cell-free release of progeny as well as cell-to-cell spread, and this tegument protein is conserved across herpesvirus subgroups (Ahlqvist and Mocarski, 2011). Virions are generally released 3-4 days after infection of the cell but can vary between strains and cell types. Production of cell-released Merlin virions for example, can first be detected at 72 hpi in fibroblasts, but not until 120 hpi in RPE-1s and U373MGs (Towler et al., 2012).

1.4.2 HCMV Latent Infection and Reactivation

All studied herpesviruses are capable of latency and re-activation, although core genes do not appear to be involved, so it is likely that this trait has been acquired in different lineages rather than through a common herpesvirus ancestor (Davison, 2011). The majority of HCMV's life cycle is spent in this 'latent'-type phase of non-productive infection, where only a small number of genes are detectable and virions are not produced (Goodrum et al., 2002, Keyes et al., 2012). HCMV genomes are also maintained in fewer cells and cell types and it is estimated that only one in every 10^5 mononuclear cells isolated from the peripheral blood are carrying HCMV latent genomes at a time (Slobedman and Mocarski,

1999). This makes latency hard to study and therefore the mechanisms of this phase in HCMV's life cycle are not currently well understood. It is also likely that HCMV does not enter 'true latency' but could instead be in a constant state of reactivation that is continually controlled by host's immune system. This has been previously described as 'smouldering persistence' (Bughio et al., 2013). HCMV infection is generally well-controlled and suppressed by the host (Sinclair and Sissons, 2006), however if the host is immunocompromised, reactivation can become a problem as viral replication can become uncontrolled. This can allow HCMV to become an opportunistic pathogen and can lead to greater morbidity and mortality (**Section 1.5**).

'Latency' or non-productive infection is important as it allows the virus to persist in the cell (and host) indefinitely. Unlike other cell types, latent HCMV-containing cells have the ability to prevent apoptosis upon viral infection which may an important factor for enabling long-term infection whilst maintaining normal functions (Wang et al., 2008). Less HCMV genes are expressed during this non-productive infection phase than in lytic infection (Goodrum et al., 2002), and these include UL133-UL138, UL144, UL82AS, US28, UL111A, IE1-x4 (a small form of IE1/UL123), a UL126a transcript, and 2 long non-coding (lnc) RNAs of 2.7 and 4.9 kb (reviewed by Goodrum (2016)). Many of these have unknown function and it is unclear what role they play in the establishment or maintenance of latency, with many dispensable for replication in fibroblasts (Dunn et al., 2003b). Most of these 'latency' genes are not solely produced during latency but are also produced during lytic infection (Dutta et al., 2015, Keyes et al., 2012). UL111A for example, encodes the viral homologue of Interleukin-10 (cmvIL-10) but alternative splicing also produces a latency-associated cmvIL-10 (LAcmvIL-10). This has been detected both *in vitro* and *in vivo* and contributes to latency-associated effects in the cell and may also function to avoid the host immune system (Sinclair, 2008, Poole et al., 2014).

HCMV 'latency' mainly occurs in undifferentiated (CD14[+]) myeloid cells and can be detected as far back as CD34[+] hematopoietic cells, the precursor to both myeloid and lymphoid progenitor cells (Goodrum et al., 2002, Maciejewski et al., 1992, Jarvis and Nelson, 2002). Other latency reservoirs are also possible, including some subsets of endothelial cells (Fish et al., 1998, Jarvis and Nelson, 2002). Although the CD34[+] precursor is a carrier of latent CMV, not all cells that have differentiated from this precursor maintain the latent virus, so primary human hematopoietic cells are often used for *in vitro* latency studies (Goodrum et al., 2002). Differentiation can play a huge role in the permissiveness of cells for infection, and once the progenitor cell has differentiated into dendritic cells (DCs) for example, HCMV

no longer remains latent in this cell type and can reactivate (Reeves et al., 2005b). It has been suggested that differentiation can cause a change in transcription factors around the MIEP region that then interact with cellular factors and subsequently modify chromatin structure to allow for transcriptional activation of IE gene expression (Reeves et al., 2005a, Sinclair and Sissons, 2006, Sinclair, 2008, Reeves et al., 2005b). This upregulation of viral immediate early genes then switches the virus into lytic phase, with IE1 or IE2 expression often used as a marker of reactivation.

1.5 HCMV Clinical disease

HCMV is ubiquitous in human populations worldwide. In the United States for example, an average of ~60% of all adults are infected, however the incidence increases to ~90% in those over 80 years old (Kenneson and Cannon, 2007, Staras et al., 2006). Seroprevalance rates around the world vary due to a range of socio-economic factors, and in some countries in the developing world, nearly 100% of individuals have acquired the virus by age 11 (Staras et al., 2006, Kenneson and Cannon, 2007). HCMV can be transmitted sexually or through a wide range of body fluids, including saliva, urine, blood and breast milk (Cannon et al., 2011). Infected young children can secrete the virus for long periods and are thus a common source of infection (Noyola et al., 2000, Cannon et al., 2011). HCMV causes a life-long infection with prolonged periods of dormancy. HCMV infection is generally asymptomatic in healthy individuals, causing only non-specific symptoms such as fever or fatigue in ~10% of individuals, which can develop 9-60 days after the primary infection (Whitley, 1996). HCMV can cause serious complications in immuno-compromised individuals, where it acts as an opportunistic pathogen. In these cases, the virus is capable of infecting the majority of organs where fever, pneumonia, retinitis, myelitis, hepatitis and other neurological and gastrointestinal issues can be associated with high levels or morbidity and mortality (Emery, 2001, Grattan et al., 1989, Lee et al., 2017, Deayton et al., 2004, Azevedo et al., 2015). Individuals undergoing organ transplantation (**Section 1.5.1**), unborn babies (**Section 1.5.2**) and HIV/AIDS patients (**Section 1.5.3**) are all particularly vulnerable to HCMV disease.

Alongside the reduction in quality and length of life for the individuals affected, HCMV also costs the health care system a vast amount of money, both for the treatment of the long term consequences of the birth defects caused by congenital HCMV and the disease outcomes of immunocompromised patients (Grosse et al., 2013). It has been estimated

that in the 1990s symptomatic congenital infection alone costs the US around $1.86 billion annually, with an average cost per child of over $300,000 (Demmler, 2006, Modlin et al., 2004). Antivirals help to reduce the incidence of disease outcomes in some patient groups (**Section 1.6.1**).

1.5.1 Transplant patients

Transplant patients are given immuno-suppressive drugs to prevent organ rejection, but this weakens the immune response which also results in HCMV infection becoming a common complication (Azevedo et al., 2015, Grattan et al., 1989). With seronegative patients, infection from a donor HCMV strain occurs in a HCMV-naïve immune system. This scenario generally holds the greatest risk for HCMV infection and disease for solid organ transplant patients (Azevedo et al., 2015, La Rosa et al., 2007, Seo and Boeckh, 2013, Emery et al., 2013). With seropositive transplant recipients, HCMV infection can instead occur either due to re-activation of an existing strain or from re-infection with a new strain from the donor. HCMV infection after transplantation increases the risk of complications, reduces the lifetime of the graft and increases the risk of organ rejection (Boppana and Britt, 2013, Azevedo et al., 2015). HCMV can also cause different complications depending on the type of transplant. For example, HCMV infection significantly increases the risk in heart transplant patients of more frequent and severe graft atherosclerosis and increased graft loss (Grattan et al., 1989). In allogenic haemopoietic (stem) cell transplant (HCT) recipients, HCMV pneumonia is a significant disease manifestation, causing over 60% of HCT deaths. It often occurs within 2 months of transplantation and comes with a poor outcome, with the median lifespan after the onset being just 25 days (Erard et al., 2015). Graft-versus-host disease (GvHD) is a situation that can occur post-transplantation and is associated with impaired graft survival and increased mortality. HCMV seropositivity is an important risk factor for acquiring GvHD, and acute GvHD has been linked to increased HCMV reactivation rates and viral loads (Broers et al., 2000).

The use of antivirals (**Section 1.6.1**) during transplantation reduces the risk of HCMV morbidity and mortality, but does not remove the risk completely, thus HCMV disease and other symptoms still occur at an unacceptable rate (La Rosa et al., 2007, Hodson et al., 2008, Erard et al., 2015, Broers et al., 2000). HCMV can also reactivate to cause disease once antivirals have been terminated.

1.5.2 HIV/AIDs patients

Individuals with HIV/AIDs are at a great risk of infections due to their compromised immune systems, and in this situation HCMV can become an opportunistic pathogen. HCMV infection can result in end-organ damage such as hepatitis, the most common being retinitis which can lead to loss of sight (Aramă et al., 2014, Casado et al., 1999, Emery, 2001). The likelihood of HCMV retinitis increases from 2% to 38% if the patient is HCMV seropositive (Casado et al., 1999). Co-infection with HCMV and detectable viremia also significantly increases the patient's progression towards AIDS and death by 2-3 times, with increased morbidity observed, even when anti-retroviral therapy was given (Aramă et al., 2014, Deayton et al., 2004, Spector et al., 1998, Adland et al., 2015). Babies who are infected with HIV also have a 3 times higher risk for acquiring a symptomatic congenital HCMV infection (**Section 1.5.3**) (Manicklal et al., 2013).

1.5.3 Congenital infection

Compared to adults, the foetal immune system is much more immature, with reduced immune responsiveness, making them more susceptible to HCMV infection (Hassan et al., 2007). The average birth prevalence of congenital HCMV is 0.64% of live births, although this can reach ~1% in some populations (Kenneson and Cannon, 2007). A large proportion of babies with HCMV infection will be asymptomatic at birth, with ~11% of them having symptomatic disease, however this varies by population and definition of 'symptomatic' (Kenneson and Cannon, 2007, Friedman and Ford-Jones, 1999). Common complications caused by congenital HCMV include damaged hearing or eyesight, hepatosplenomegaly, jaundice, neurological damage and even death (Fowler et al., 1992, Friedman and Ford-Jones, 1999, Boppana et al., 2005, Pass et al., 2006). The high frequency of these serious complications causes HCMV to be the leading infectious cause of birth defects (Luo et al., 2010). HCMV is also the leading cause of non-hereditary sensorineural hearing loss (SNHL) in children, with symptomatic children at a much higher risk (Friedman and Ford-Jones, 1999, Ogawa et al., 2007, Boppana et al., 2005). The majority of symptomatic children develop SNHL, brain disorders or behavioural problems, along with up to 15% of those who were asymptomatic at birth (Luo et al., 2010, Pass et al., 1980, Fowler et al., 1993, Cheeran et al., 2009).

The rate of transmission to the foetus increases during the course of gestation, however neurological outcomes are more severe when infection occurs during the first trimester (Pass et al., 2006, Enders et al., 2011). Seronegative women who undergo a primary HCMV infection are more likely to have a child with more serious congenital HCMV infection

(Fowler et al., 2003, Fowler et al., 1992). Although likely due to the increased likelihood of primary infection, maternal age can also be a risk factor for congenital HCMV, with foetuses of younger women (under the age of 25), more likely to have more severe infection and symptoms, (Fowler et al., 2003, Fowler et al., 1993, Kenneson and Cannon, 2007, Lanzieri et al., 2014). Despite this, recurrent infections are more common in general due to the large number of seropositive women, so recurrent infections are believed to account for the majority of the overall disease burden, especially in low-income populations (Cheeran et al., 2009, Stagno et al., 1986, Ross et al., 2006, Wang et al., 2011). Many other risk factors are involved which often overlap or impact on each other, so it can be hard to pinpoint which are significant or identify which women are most at risk during their pregnancy (Kenneson and Cannon, 2007, Staras et al., 2006, Fowler et al., 1993).

1.5.4 Other at-risk groups

In conjunction with the common patient groups above, and those with autoimmune diseases (Lee et al., 2017), HCMV has also been proposed to play a role in chronic inflammatory diseases, vascular disease, atherosclerosis, and accelerated immune senescence in the elderly (reviewed by Boppana and Britt (2013)). It is likely that the host inflammatory responses plays a big part in these outcomes, but overall, HCMV seropositivity is a major risk factor, and high viral loads have a huge impact on the severity of infections and disease outcomes (Boppana et al., 2005, Aramă et al., 2014, Lee et al., 2017, Broers et al., 2000, Spector et al., 1998). HCMV DNA load can therefore be a useful predictor of at-risk patients (Spector et al., 1998, Boppana et al., 2005, Deayton et al., 2004, Lee et al., 2017).

1.6 HCMV Antiviral Treatments and Vaccines

1.6.1 Antiviral treatments

Those who are immunosuppressed can be given antivirals to help reduce the risk and severity of primary HCMV infection. Some of the antivirals currently given include Ganciclovir, Valganciclovir, Foscarnet and Cidofovir (Cheeran et al., 2009). All work by targeting DNA replication but have different side effects and are preferentially given in different cases of HCMV infection. Ganciclovir is an example of a nucleotide analogue, which can competitively inhibit the cellular nucleotides to stop replication, and can cause DNA chain termination (Chen et al., 2014). Once Ganciclovir is phosphorylated by HCMV's pUL97 and then by cellular kinases, it mimics the cellular guanine nucleoside dGTP and

competitively inhibits its incorporation into DNA. This can preferentially inhibit the viral DNA polymerase over cellular DNA polymerases in order to preferentially disrupt viral DNA synthesis (Chen et al., 2014). Ganciclovir was the first compound licensed specifically for the treatment of severe HCMV infections and it is usually given to bone marrow, solid organ transplant recipients or HIV patients (Cheeran et al., 2009, Hodson et al., 2008). Ganciclovir can reduce the risk of infection and disease, as well as increasing survival rates in some patient groups, although this was not always by a significant amount (Hodson et al., 2008, Lee et al., 2017, Nakamae et al., 2011). It does also have some side effects, and is prone to causing a reduction of mature blood cells (cytopenia) which occurs in nearly 40% of HCT patients, although this is generally reversible (Nakamae et al., 2011). Neurological dysfunction can also occur, which was more common with Ganciclovir than Valganciclovir in transplant patients (Hodson et al., 2008). Valganciclovir is an ester of Ganciclovir and is rapidly metabolized into Ganciclovir following oral administration. It has similar side effects and efficacy as IV Ganciclovir but is better absorbed than oral Ganciclovir, even in patients with GvHD (Einsele et al., 2006, Hodson et al., 2008). It is generally used for treatment of HIV patients with HCMV retinitis, as well as for prevention of HCMV infection in transplant patients, especially those at high risk (Einsele et al., 2006).

Both Foscarnet and Cidofovir have similar mechanisms in which they inhibit the HCMV DNA polymerase; with Cidofovir requiring phosphorylation by cellular kinases to be selectively incorporated into the viral DNA chain, whereas Forscarnet does not require phosphorylation to competitively inhibit the polymerase (Cheeran et al., 2009, Seo and Boeckh, 2013). Treatment times with Cidofovir can be more flexible, due to its long intracellular half-life but both are second-line treatments due to their toxicity effects. Nephrotoxicity is the most significant of these side effects, and necessitates pre-hydration, requiring intravenous (IV) administration and added cost (Bregante et al., 2000, Einsele et al., 2006, Ljungman et al., 2001). Side effects such as electrolyte abnormalities, neutropenia, diarrhoea, liver toxicity and metabolic acidosis can also occur, however with Foscarnet there are no risks of haematological toxicity and is therefore recommended for patients with bone marrow failure (Bregante et al., 2000).

Maribavir is one of the newer antivirals, which works by inhibiting the UL97 kinase and preventing viral encapsidation and nuclear egress. It has only mild toxicity effects such as taste disturbance and resulted in the clearance of HCMV DNA in patients who had

resistance to other treatments (Avery et al., 2010). Other potential new drugs include Leflunomide (a rheumatoid arthritis drug), Artesunate (an anti-malaria drug), and Letermovir (a new CMV-specific drug). HCMV immunoglobulin (Ig) has also been used, often in combination with one of the main current antivirals (Avery et al., 2010, Seo and Boeckh, 2013, Cheeran et al., 2009, Ljungman et al., 1992, Nigro et al., 2005, Marty et al., 2017). HIV/AIDS patients are currently given highly active antiretroviral therapy (HAART) which reduces the risk of HCMV related death, as well as retinitis, although unexpectedly can instead result in inflammatory vitritis which could also lead to loss of sight (Deayton et al., 2004, Karavellas et al., 1999).

Due to a relatively small selection of antivirals, problems with resistance can occur, including Ganciclovir resistance which is primarily caused by mutations in the UL97 kinase gene. This has especially become a problem for HIV patients, pancreas, or kidney and pancreas transplant patients, with many Ganciclovir-resistant isolates also resistant to Cidofovir (Limaye et al., 2000). In one study, Ganciclovir-resistant HCMV disease made up 2.1% of the overall HCMV occurrence, with antiviral-resistance a cause of morbidity and mortality, especially in HCT recipients and organ transplant recipients with seropositive donors (Avery et al., 2010, Limaye et al., 2000). Mutations in the viral DNA polymerase gene (UL54) can confer resistance to Foscarnet (Limaye et al., 2000, Cheeran et al., 2009), and this is becoming more common due to the increased use of antivirals (Chen et al., 2014, Emery, 2001). It is especially problematic if a patient has mutations in both UL54 and UL97 genes as it severely limits treatment options.

Antivirals and HAART therapy have reduced the incidence of disease outcomes in some patient groups, but the levels of morbidity and mortality are still high across a wide range of people due to HCMV infection (Deayton et al., 2004, Emery, 2001, Erard et al., 2015, Broers et al., 2000). HCMV is therefore still considered a high priority vaccine target by the National Vaccine Advisory Board (Modlin et al., 2004).

1.6.2 Vaccines

While HCMV antivirals are clearly extremely important, there remains an urgent need to generate an effective preventative vaccine, particularly to prevent congenital infections. To date, vaccines against HCMV have been unsuccessful or have limited efficacy. Main strategies adopted include attenuated or modified versions of HCMV, individual antigen vaccines, or virus-like particles (**Table 1.4)**.

Table 1.4: A summary of different HCMV vaccine strategies that have been developed and trialled, with examples of each type.

Type of vaccine	Examples
Live virus (attenuated HCMV)	Attenuated AD169 (Elek and Stern, 1974)
	Attenuated Towne (Plotkin et al., 1994, Plotkin et al., 1984, Plotkin et al., 1991)
	Towne–Toledo recombinant chimeras (Adler et al., 2016)
Viral vectors	Canarypox vector expressing gB (Adler et al., 1999, Bernstein et al., 2002)
	Alphavirus replicon expressing HCMV gB, pp65 and IE1 (Bernstein et al., 2009)
	Modified vaccinia ankara (MVA) virus expressing all five proteins of the gH/gL pentameric complex (Wussow et al., 2014)*
Defined antigens/ peptides/ recombinant proteins	Glycoprotein B (gB) with MF59 adjuvant (Pass et al., 2009, Pass et al., 1999, Griffiths et al., 2011)
	Peptide of pp65 and tetanus epitopes, with TLR9 agonist as an adjuvant (Nakamura et al., 2016).
	Pentameric complex (Genini et al., 2011, Gerna et al., 2017, Lilleri et al., 2013)
DNA	DNA vaccine of gB and pp65 with adjuvant CRL1005 and benzalkonium chloride (Kharfan-Dabaja et al., Smith et al., 2013a)
Virus like particles	Non-infectious dense bodies (from AD169 and Towne) (Cayatte et al., 2013)

*Not yet in human studies

1.6.2.1 Attenuated HCMV vaccines

HCMV vaccine development in the 1970s used live attenuated viruses such as the AD169 strain (Elek and Stern, 1974) or the Towne strain that had been extensively passaged in fibroblasts to diminish its pathogenicity (Plotkin et al., 1975). The attenuated Towne strain was deemed safe in healthy volunteers, but mothers of infected children were not protected from acquiring HCMV (Adler et al., 1995). It also did not significantly affect the rate of infection in kidney transplant patients who had HCMV seropositive donors, although it did protect against more severe HCMV disease outcomes (Plotkin et al., 1984, Plotkin et al., 1994, Plotkin et al., 1991). It was considered that the Towne virus may be too attenuated, so chimera viruses of Towne and Toledo were made and were tested in a phase 1 study. These chimera viruses did not cause disease, were not shed by vaccinated individuals, and the majority of patients produced neutralising activity and CD8+ T-cell responses, both of which could be detected a year after vaccination occurred (Adler et al., 2016). Attenuated live viruses are valuable due to their similarity to wildtype strains, but proving their safety can be challenging and there is a small chance that they could potentially be modified and reactivate in the patient (Anderholm et al., 2016).

1.6.2.2 Subunit antigen vaccines

A simpler strategy is to generate a subunit vaccine based on key HCMV antigens. Such vaccines have been designed with or without adjuvants and generally focus on the envelope glycoprotein B (gB) (Frey et al., 1999, Griffiths et al., 2011, Pass et al., 1999, Pass et al., 2009). In a phase 2 trial, a gB vaccine with an MF59 adjuvant was found to reduce the duration of viraemia and thus reduce the number of days on pre- emptive antiviral therapy, with gB antibody titres remaining high as the patients progressed to transplantation (Griffiths et al., 2011). Moreover, women of childbearing age who received the vaccine were twice as likely to remain uninfected from HCMV, although conclusions about incidence of congenital infection could not be formed (Pass et al., 2009). Other subunit vaccine approaches have involved immunogenic targets such as the tegument protein pp65 (UL83) or proteins in pentameric complex gH/gL/UL128-131 (Genini et al., 2011, Gerna et al., 2017, Lilleri et al., 2013, Nakamura et al., 2016).

1.6.2.3 Virus like particles

Another strategy is vaccines that express some of these immunogenic proteins from viral vectors such as a canarypox vector (Adler et al., 1999, Bernstein et al., 2002), a modified vaccinia ankara (MVA) virus (Wussow et al., 2014), or an alphavirus replicon particles

vaccine based on Venezuelan equine encephalitis virus (VEEV) (Bernstein et al., 2009, Loomis et al., 2013)**.** These viral vectors are non-replicating but can provide high expression of proteins and can boost both the cellular and humoral immune responses to the presented peptides (Bernstein et al., 2009). Any strategy that can improve or increase the involvement of the host's immune system are potentially beneficial to a vaccine. One vaccine named CMVPepVax has been trialled that uses a chimeric peptide that consists of a cytotoxic CD8 T-cell epitope from pp65 and a tetanus T-helper cell epitope, alongside a Toll-like receptor 9 (TLR9) agonist as an adjuvant to drive cellular immunity (Nakamura et al., 2016). In HCT patients the vaccine caused a significant increase in pp65 specific T cells and provided an increased protection from HCMV viraemia and reactivation as well as the reduction of adverse events, and a reduction of the usage of antivirals. The only downside to this vaccine is that it can only be given to 40% of the population who are HLA A*0201, but it shows a promising step in the right direction (Nakamura et al., 2016).

Other vaccine strategies have used viral DNA (Kharfan-Dabaja et al., 2012, Smith et al., 2013a), or dense bodies (Cayatte et al., 2013). Overall, several vaccines that have been trialled have had an impact in lessening the severity of HCMV symptoms but often do not affect the rate of infection, and do not completely eliminate the problem of HCMV disease or mortality as much as hoped. Other treatments trialled have included the use of antibodies and immunoglobulin/hyperimmune globulin (Ishida et al., 2017, Nigro et al., 2005, Genini et al., 2011, Lilleri et al., 2013).

1.7 Immune Responses to HCMV

The human immune system is highly complex and consists of 2 main branches- innate immunity (**Section 1.7.1**) and adaptive immunity (**Section 1.7.2**). There is also a third type of immunity which is mechanistically distinct from innate and adaptive immunity. It is similar to adaptive immunity in that it targets specific viruses, however it does not respond differently upon subsequent infection with the same pathogen. Intrinsic immune proteins are constitutively expressed and can therefore act against the virus immediately after infection (Rossini et al., 2012, Saffert and Kalejta, 2006). One example is the human death-domain associated protein (hDaxx) which silences viral promoters (Saffert and Kalejta, 2006), however this type of immunity will not be covered in this thesis.

1.7.1 Innate immunity

The innate branch of the immune system is a constant detection system that is non-specific to the pathogen or environment that has triggered it. This includes physical barriers such as the skin, non-specific antimicrobial peptides, and includes the processes of inflammation and phagocytosis to non-specifically kill pathogens (Pfeffer, 2003, Boehme et al., 2006, Boehme et al., 2004, Thompson et al., 2011). Host cells have many pathways that can be triggered in response to pathogens such as HCMV, with HCMV triggering a similar response to the one generated by interferon (IFN) (Hertel and Mocarski, 2004, Simmen et al., 2001). The principle cell types and processes involved are detailed below.

1.7.1.1 Pattern recognition receptors

Cells can detect microbe-, pathogen- or damage-associated molecular patterns (MAMPS/PAMPS/DAMPs), which can be cellular proteins that come from damage or abnormalities, cellular DNA or extracellular matrix proteins, or the presence of microbial or viral proteins and nucleic acids including viral DNA (Thompson et al., 2011, DeFilippis et al., 2010, Boehme et al., 2006, E and Kowalik, 2014). These molecular patterns are recognised through pattern recognition receptors (PRRs), including Toll-like receptors (TLRs) and nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs), that can respond quickly and non-specifically to infections such as HCMV (Muralidharan and Mandrekar, 2013, Thompson et al., 2011, Rossini et al., 2012). Different cell types contain different receptors which can have different signalling outputs, as well as different responsiveness to different ligands (Barr et al., 2007, Kaisho and Akira, 2006).

Activation of PRRs can lead to downstream inflammatory signalling which triggers the activation of transcription factors including the nuclear factor kappa-B (NF-κB), mitogen-activated protein kinase (MAPK) and interferon regulatory factors (IRFs) 1, 3, 5 and 7 (Arango Duque and Descoteaux, 2014, Deng et al., 2004, Thompson et al., 2011, Muralidharan and Mandrekar, 2013). For example, the binding of HCMV gB and gH to TLR-2 upon viral entry triggers multiple signalling pathways (**Section 1.4.1.2**), and the detection of HCMV double-stranded DNA by the PRR Z-DNA binding protein 1 (ZBP1) is essential for IRF3 activation and IFN-β expression (DeFilippis et al., 2010)**.**

1.7.1.2 Apoptosis

Apoptosis is highly regulated programmed cell death, and it can be triggered in multiple ways, both intrinsically or extrinsically. Intrinsic triggering of apoptosis can occur when DNA damage is detected but cannot be repaired (E and Kowalik, 2014), or when

unfolded/denatured proteins cannot be re-folded (Isler et al., 2005), and these stresses can be mediated by viruses such as HCMV. Intracellular stresses can often lead to the release of cytochrome c from mitochondria which results in the formation of a protein complex called the apoptosome, which subsequently activates the caspase cascade, leading to apoptosis (Kominami et al., 2012, Peter and Krammer, 2003).

Apoptosis can also be triggered extrinsically where the process is initiated by signals from cytotoxic effector cells such as cytotoxic T lymphocytes (CTLs) or NK cells. They can trigger apoptosis extrinsically through the release of perforin which forms pores in the infected cell's membrane, allowing granzymes (proteolytic enzymes) to enter. This activates the downstream cellular pathways that lead to apoptosis (Smyth et al., 2001, Lieberman, 2010, Pegram et al., 2011, Barber, 2001). Another method is through the activation of 'death receptors' such as CD95/Fas which leads to caspase activation (Seirafian et al., 2014, Peter and Krammer, 2003, Kominami et al., 2012, Goldmacher, 2005). Once the process of apoptosis has been triggered it cannot be stopped, with effector caspases cleaving host proteins, causing the cell to be broken down into cell fragments called apoptotic bodies. These are then phagocytosed by neighbouring cells such as by macrophages, destroying both the cell and the internal pathogen, limiting viral spread (Peter and Krammer, 2003, Holcik and Sonenberg, 2005, García-García and Rosales, 2000-2013).

1.7.1.3 Cytokine and chemokine responses

When immune cells are triggered they can activate the production of type I and 2 interferons (IFN), pro-inflammatory cytokines (such as IL-1, IL-12 and TNF-α) and chemokines (Muralidharan and Mandrekar, 2013, Khairallah et al., 2017, Varani et al., 2007). Inflammatory chemokines can attract cells from the bloodstream towards the site of infection (Megjugorac et al., 2004, Khairallah et al., 2017, Thompson et al., 2011, Arango Duque and Descoteaux, 2014). For example, upon contact with HCMV, plasmacytoid dendritic cells (pDCs) produce CCL3 (a β-chemokine) which attracts NK cells, Th1 cells, B cells and immature myeloid DCs to sites of viral infection (Megjugorac et al., 2004, Varani et al., 2007). Neutrophils, NK cells and CTLs are recruited rapidly, with cytokines such as IL-12 and IL-18 able to enhance their cytolytic activity (Watford et al., 2003, Biron et al., 1999). Cytokines can prevent the replication of a variety of viruses and also increase the sensitivity of neighbouring cells, sensitizing them to apoptosis which may inhibit the spread of infection (Barber, 2001, Muralidharan and Mandrekar, 2013). Cytokines also lead to the activation and differentiation of immune cells (Griffin et al., 2012, Watford et al., 2003, Arango Duque and Descoteaux, 2014, Pfeffer, 2003, Katze et al., 2002), including dendritic

cells (DCs) (Megjugorac et al., 2004) and NK cells (Biron et al., 1999, Nguyen et al., 2002). This is also true for cells of the adaptive response such as T and B lymphocytes (Ben-Sasson et al., 2009).

1.7.1.4 Antigen presenting cells (APCs)

Professional antigen-presenting cells (APCs) such as macrophages and DCs play a huge role in the crossover between the innate and adaptive branches of the immune system. APCs can intracellularly degrade pathogenic proteins such as by phagocytosis where the pathogen is subject to lysosomal digestion by enzymes and the production of reactive oxygen and nitrogen species (ROS, NOS) (Nathan and Shiloh, 2000, Muralidharan and Mandrekar, 2013). HCMV antigens can then be presented on cell surface major histocompatibility complexes (MHC) to T cells. With HCMV infection it is often peptides from abundant proteins such as pp65 and IE1 that are presented to CTLs on MHC-I, and MHC upregulation by the host is one way to attempt to increase the recognition by T cells (Besold et al., 2007, McLaughlin-Taylor et al., 1994, Biron et al., 1999). Most APCs will also be stimulated to produce T-helper1 (Th1)-inducing cytokines, such as IL-12 and IL-18; with pDCs producing unusually high amounts of type I IFNs (Megjugorac et al., 2004, Varani et al., 2007, Thompson et al., 2011, Kaisho and Akira, 2006, Smith et al., 2008).

1.7.1.5 Natural killer (NK) cells

Classical NK cells play a role in clearing virus-infected cells and receive activating and inhibitory signals (ligands) from a broad range of cell surface receptors, and it is the balance of these signals that determines whether or not the NK cell is activated to kill the target cell. The main families of NK receptors are the killer immunoglobulin-like receptors (KIR) family, leukocyte immunoglobulin-like receptors (LIRS), the natural cytotoxicity receptor (NCR) family and the NKG2 (CD94) family. The KIRs, LIRs and NCRs belong to the immunoglobulin (Ig) superfamily, and the NKG2 family are C-lectin type receptors (Pegram et al., 2011). Each family tends to contain both activating and inhibitory receptors, for example in the NKG2 family, NKG2A and B are inhibitory receptors and NKG2C, D and E are activating receptors (La Rosa and Diamond, 2012, Prod'homme et al., 2007). The main inhibitory ligands on target cells are endogenous major histocompatibility complex class I (MHC-I) molecules. NK cells are also unique in that they can detect reduced or atypical MHC-I molecules, termed "missing self" (Pegram et al., 2011, Gasser and Raulet, 2006, Ljunggren and Karre, 1986).

The NKG2D activating receptor is unusual in recognizing multiple ligands that have been shown to be upregulated when cells are stressed (e.g. by virus infection) (Eagle and Trowsdale, 2007, Rolle et al., 2003, Eagle et al., 2009). These ligands include the MHC class I-polypeptide related sequences (MIC) A and B, and the human UL16-binding protein (ULBP) proteins 1-6; and their increase can potentially tip the balance and activate the NK cell (Prod'homme et al., 2007, Rolle et al., 2003, Eagle et al., 2009, Slavuljica et al., 2011). NKG2D is expressed not only on NK cells, but also on CD8$^+$ T cells, some γδ T cells, some NK-T cells, and activated macrophages (Slavuljica et al., 2011, Wilkinson et al., 2008).

Activating receptors can signal through immuno-receptor tyrosine-based activating motifs (ITAMs) using DAP-12, or through phosphatidylinositol-3 kinase (PI3K) and other pathways via DAP-10; whereas inhibitory receptors signal through intracellular immuno-receptor tyrosine-based inhibitory motifs (ITIMs) instead (Wu et al., 1999, Lanier et al., 1998, Pegram et al., 2011). NKG2D associates with DAP10 in order to produce a strong activating signal (Lanier, 2008, Wu et al., 1999).

As mentioned, it is the balance of inhibitory and activating signals that determines NK activation. If activated, the NK cell stimulates cytotoxicity of the infected target cell, mediated by the release of perforin and granzymes that lead to apoptosis (**Section 1.7.1.2**). This can also be brought about by Fc receptor (CD16) ligation with Ig molecules bound to cell surface antigens on the target cell, mediating killing by antibody dependent cell-mediated cytotoxicity (ADCC) (reviewed by Aicheler et al. (2013) and Biron et al. (1999)). NK cells can also kill target cells via the ligation of death receptors such as Fas with their ligands, their activation also leading to apoptosis (**Section 1.7.1.2**) (Seirafian et al., 2014, Peter and Krammer, 2003). NK activation can also lead to the production of cytokines and chemokines (**Section 1.7.1.3**) which can attract and activate cells of the adaptive response (Wu et al., 1999, Lanier, 2008, La Rosa and Diamond, 2012, Biron et al., 1999). Cytokines such as IFN, IL-12 and IL-15 can stimulate NK cells and can regulate NK cell responses, inducing NK cytotoxicity, NK IFN-gamma expression, NK cell accumulation and proliferation, and protective NK cell responses (Nguyen et al., 2002, Biron et al., 1999).

Due to all of these early detection systems of the innate immunity, host cells can activate anti-viral immune responses before the outset of viral replication (Boehme, Singh, Perry, & Compton, 2004).

1.7.2 Adaptive response

The adaptive arm of the immune response takes longer to take effect due to the complicated activation process that can take up to 7 days. It can however provide a pathogen-specific response and can lead to lasting memory that provides a quicker response to a future attack (La Rosa and Diamond, 2012). Cells of the adaptive immune system recognising pathogen-specific antigens and distinct epitopes instead of PAMPs or DAMPs. This process involves both B and T lymphocytes which can recognise antigen epitopes directly, or through antigen peptide presentation from MHCs on APCs (**Section 1.7.1.4**) respectively. The receptors on T and B lymphocytes (TCRs and BCRs/ immunoglobulins respectively) are generated through gene rearrangements, providing them with a highly varied array of receptor combinations allowing an unlimited range of detection for potential pathogen epitopes, with a system in place that selects away from the detection of self-antigens (Janeway et al., 2001, Spits et al., 1995). Once activated by an antigen, both T and B lymphocytes can clonally replicate within ~96 hours (Brinkworth and Thorn, 2013).

1.7.2.1 B lymphocytes

B cells express receptor immunoglobulins (Igs) or antibodies that are made of heavy and light chains, that are found membrane-bound as well as in secreted forms. These are often directed against abundant HCMV proteins such as pp65, pp150, gB, gH, and gH/gL multimeric complexes as well as non-structural proteins such as IE1 (La Rosa and Diamond, 2012, Dauby et al., 2014). B cells are activated and differentiated into Ig-secreting plasma cells with cytokines (**Section 1.7.1.3**), often need the presence of T cells or T cell-secreted IL-2 for antibody production (Varani et al., 2007). It is around day 7 that plasma cells are detected, but longer for antibody secretion to be observed (Varani et al., 2007). Antibodies can sometimes neutralise the pathogen directly e.g. by blocking a binding receptor as exemplified by gB antibodies preventing HCMV entry into cells (Gerna et al., 2008). Antibodies can also mark the infected cell for attack by other immune cells (such as NK cells) through their Ig receptors, with an outcome of antibody dependent cell-mediated cytotoxicity (ADCC) (Aicheler et al., 2013).

Upon activation, B cells can also make cytokines such as IL-10 and IFN-γ so have the potential to affect T cell differentiation (Barr et al., 2007). Primary HCMV infection also mobilizes a large pool of memory B cells (MBCs) that includes activated and atypical MBCs (Dauby et al., 2014).

1.7.2.2 T lymphocytes

T cell receptors (TCRs) are made up of either an α and β chain, or a γ and δ chain, making the 2 distinct lineages αβ and γδ T cells, with T cell recognition occurring via the MHC class I and II antigen presentation pathways displayed from APCs (**Section 1.7.1.4**) (Janeway et al., 2001). Over the course of infection, T cells populations change and differentiate between naïve, effector and memory T cells, and it is the αβ T cells that develop into CD4+ and CD8+ T effector cells, with helper T cells (Th) expressing CD4, and cytolytic T lymphocytes (CTLs) expressing CD8 (Ben-Sasson et al., 2009, Whitmire, 2011, Sylwester et al., 2005, Janeway et al., 2001). HCMV-specific CD4+ T cells will start to circulate ~7 days after viral replication starts, with HCMV-specific CD8+ T cells appearing in the peripheral blood later on (La Rosa and Diamond, 2012). Subtype differentiation is triggered by cytokine responses such as IL-12 (Watford et al., 2003, Varani et al., 2007, Barr et al., 2007, Ben-Sasson et al., 2009). Subsets themselves also produce different cytokines and have different effects, including upregulation of MHC presentation and aiding B cell differentiation (Biron et al., 1999, La Rosa and Diamond, 2012, Whitmire, 2011). CD8+ CTLs are able to lyse HCMV-infected cells using perforin and granzyme B as described above (Section 1.7.1.2) (La Rosa and Diamond, 2012, Lieberman, 2010).

Memory CD8+ cells form later on, upon recovery from HCMV infection, and provide important long-term protection against the antigen and recurrent infections, with protective immunity shown to be transferred through adoptive therapy of HCMV-specific T cells (La Rosa and Diamond, 2012, Besold et al., 2007). Memory T cells retain their cytolytic potential and can be induced to expand vigorously and rapidly (La Rosa and Diamond, 2012, Dauby et al., 2014, Khairallah et al., 2017). The CD4+ and CD8+ T cell responses are mainly directed against pp65 and IE1 proteins, although can also be against other tegument proteins such as pp28, capsid proteins, or envelope proteins such as gB (Besold et al., 2007, McLaughlin-Taylor et al., 1994, Sylwester et al., 2005).

The T cell response raised against HCMV is larger than that to any other virus, with HCMV-specific memory T cells on average making up ~10% of both the CD4+ and CD8+ memory compartments in peripheral blood (Sylwester et al., 2005). The activation of both memory CD4+ T cells and CD8+ T cells are required for a more protective CTL response to recurrent infection (Whitmire, 2011).

1.7.3 Summary

The immune system is extremely complex, and all parts of the system overlap, with different cells signalling to each other and affecting each other's activation and differentiation. These cell types are all important and they work together to achieve the desired outcome of preventing viral replication. Individuals who are missing parts of this defence system have a compromised immune system and can be more susceptible to diseases or have increased disease severity such as in the clinical disease situations covered above (**Section 1.6**). This is particularly true for individuals with HCMV who have a defect in their NK responses (reviewed by Orange (2013).

1.7.4 Immune Modulation by HCMV

HCMV has evolved along with its host to encode a remarkable variety of ways to avoid or modulate the immune system. Firstly, HCMV can cause transcriptional changes within the host cell, and regulate host cell proteins. Transcriptional changes can be seen throughout the course of infection, starting from ~15 mins post infection, with viral protein synthesis overtaking cellular protein synthesis by 48 hpi (Hertel and Mocarski, 2004, Simmen et al., 2001, Stinski, 1977). HCMV targets host genes that are involved in regulation of the cell cycle, DNA replication, cell surface receptors, energy production and inflammation, including a high degree of modulation of interferon stimulated genes (Hertel and Mocarski, 2004, Simmen et al., 2001). A main player in this role is the abundant tegument protein pp65 (ppUL83), which blocks the induction of some of the IFN-responsive genes, restricting the activation of NF-κB and IRF1 and blocking IFN-α signalling, making the cells less responsive to endogenous cytokines (Browne and Shenk, 2003). pUL144 (a member of the TNF receptor family) that can activate the NF-κB pathway, enhancing the expression of the chemokine CCL22, which can attract Th2 and regulatory T cells and may help with immune evasion (Poole et al., 2006).

1.7.4.1 T cell evasion

Another important immune evasion strategy is to reduce the detection of infected cells, principally through the down-regulation of antigen presentation by MHC-I and II to reduce T cell (**Section 1.7.2.2**) activation, as well as reducing the presentation of some of the NK activating ligands (**Section 1.7.3.2**), reducing the infected cell's susceptibility for cytotoxic killing. Mechanisms involved in this process include MHC retention in the ER, increased degradation, and prevention of the transport of peptides via the TAP transporter (Besold et al., 2007; Hsu et al., 2015; Noriega et al., 2012).

US2 and US6 family proteins, along with pp71 (ppUL82) are all involved in modulating MHC-I molecules (**Section 1.8.7, 1.8.11 and 1.8.12**) (Hsu et al., 2015, Huber et al., 2002, Noriega et al., 2012b, Pande et al., 2005, Besold et al., 2007, Furman et al., 2002, Tirabassi and Ploegh, 2002, Trgovcich et al., 2006). pUS2 and pUS3 also play a role in MHC-II reduction and pUS2 also degrades an additional array of binding ligands, reducing integrin signalling, cell adhesion and cell migration (**Section 1.8.11**) (Hsu et al., 2015, Tomazin et al., 1999).

1.7.4.2 NK cell evasion

Reductions in MHC presentation will impact NK cells which detect MHC ligands through activating and inhibitory receptors (**Section 1.7.1.5**). HCMV encodes an array of genes that are capable of suppressing NK cell recognition- UL16, UL18, UL40, UL83, UL135, UL141, UL142, US18 and US20, along with microRNA (miR-UL112) (Wilkinson et al., 2008, Wilkinson et al., 2015). The overall consequence is to effectively alter the balance of the NK activating and inhibiting receptors signals to tip the balance towards non-activation of the NK cell towards the infected cell, to allow the infected cell to persist undetected. Multiple HCMV genes function by downregulating stress-induced ligands for NK activating receptors, including UL16, miRUL112, UL141, US18 and US20. gpUL16 for example reduces the cell surface of three NKG2D ligands, MICB, ULBP1 and ULBP2, by retaining them intracellularly (Welte et al., 2003, Cosman et al., 2001, Sutherland et al., 2002). HCMV miRUL112 also downregulates MICB by preventing translation of the MICB mRNA (Stern-Ginossar et al., 2007). pUS18 and pUS20 target both the NKG2D ligand MICA, and NKp30 ligand B7-H6 for degradation (**Section 1.8.15**) (Fielding et al., 2014, Fielding et al., 2017). HCMV gpUL141 and gpUL142 also provide resistance to NK attack (**Section 1.8.3 and 1.8.4**. These 2 genes are located within the UL/*b'* region and helps explain why some laboratory strains of HCMV that are missing this region are less pathogenic and more easily targeted for cytolysis (Wilkinson et al., 2015, Wilkinson et al., 2008, Cha et al., 1996).

HCMV signal peptide (sp)UL40 doesn't reduce an NK activating ligand, but instead upregulates the non-classical MHC molecule HLA-E, which binds to the NK inhibitory receptor CD94/NKG2A (Prod'homme et al., 2012, Tomasec et al., 2000, Wang et al., 2002). Interestingly, The HLA-E binding peptide derived from spUL40 upregulates cell surface expression of gpUL18 (Prod'homme et al., 2012). HCMV gpUL18 encodes an MHC class I homolog binds the inhibitory leukocyte Ig-like receptor 1 (LIR1) which is found on NK cells as well as B lymphocytes, monocytes, dendritic cells, and subsets of T cells (Prod'homme et al., 2007). gpUL18 strongly inhibits NK cells by binding to LIR1 (**Section 1.8.4**) (Prod'homme

et al., 2007), although its role is a complicated and controversial one, as although cells containing the LIR1 receptor were inhibited, another NK subset which were LIR1 negative were activated (Prod'homme et al., 2007).

The viral tegument protein pp65 (ppUL83) has a different mechanism for NK evasion which involves targeting the activating receptor NKp30 itself, causing dissociation of the linked CD3ζ chain and leading to general inhibition of cytotoxicity against the infected cells (Arnon et al., 2005). Overall, HCMV manages to balance the down-regulation of MHC-I molecules without alerting the host to 'missing self' by modulating ligands for both activating and inhibitory receptors.

1.7.4.3 Other immune modulation functions

Other than HCMV's multiple mechanisms for the evasion of NK cell and T cell recognition and killing, it also encodes other functions for mechanisms to avoid apoptosis through the inhibition of death receptors, through Fc binding proteins and homologues of cellular cytokines or receptors. For example, HCMV encodes 3 proteins that can target death receptors to protect the infected cell against both soluble TRAIL and TRAIL-dependent killing. gpUL141, aside from blocking CD155, can also bind directly to TRAIL receptor 2 (TR2) and sequester this death receptor in the endoplasmic reticulum (Smith et al., 2013b). HCMV pUL35 and pUL37 are instead involved in the downregulation of Fas (CD95) to prevent Fas-mediated apoptosis of the cell (Seirafian et al., 2014). pUL36 achieves this by being a viral inhibitor of caspase-8 activation (vICA) (**Section 1.8.13**), whilst pUL37 is a viral mitochondrion-localised inhibitor of apoptosis (vMIA) that inhibits apoptosis by suppressing cytochrome c release (**Section 1.7.1.2**) (Skaletskaya et al., 2001, Goldmacher, 2005, Seirafian et al., 2014). pIE1, pIE2 and pUL38 have also been suggested to function as inhibitors of apoptosis (Zhu et al., 1995, Terhune et al., 2007).

Human cytomegalovirus (HCMV) is unique among viruses in that it encodes an array of proteins that recognize the constant Fc domain of IgG, called Fcγ-binding receptors (FcγRs), that are found on the cell surface and function as rivals of host FcγRs (Corrales-Aguilar et al., 2014a). These viral FcγRs (RL11, RL12, RL13, and UL119-118) mostly belong to the RL11 multigene family (**Section 1.8.2**) (Corrales-Aguilar et al., 2014a, Cortese et al., 2012). Both gpRL11 (gp34) and gpUL119-118 (gp68) have shown to have a broad inhibition effect towards 3 host Fc receptors (FcγRIIIA, FcγRIIA and FcγRI) which can reduce or prevent the antibody-mediated triggering of antiviral immunity (Corrales-Aguilar et al., 2014b).

Other HCMV mimics of host receptors includes chemokine receptor homologues pUS28, pUL33 and pUL78 (**Section 1.8.14**) (Rosenkilde et al., 2001, Tadagaki et al., 2012, Bodaghi et al., 1998), and UL22A which is a viral chemokine receptor decoy that binds to the chemokine RANTES to block its interaction with cellular receptors (Wang et al., 2004). pUL146 and pUL147 on the other hand, are chemokine homologues (**Section 1.8.9**) (Scarborough et al., 2017, Penfold et al., 1999, Lüttichau, 2010). HCMV also encodes a homologue of human IL-10, pUL111A, termed viral IL-10 (vIL-10,) (Kotenko et al., 2000, Chang et al., 2004). This has low sequence similarity to human IL-10 but is biologically active and still has immunosuppressive properties (Spencer et al., 2002, Kotenko et al., 2000).

Additional HCMV modulation proteins include the antiviral protein kinase R (PKR) antagonists pTRS1 and pIRS1 (**Section 1.8.13**) (Ziehr et al., 2016, Colberg-Poley, 1996, Child et al., 2004), and pUL135 which inhibits synapse formation by remodelling the actin cytoskeleton to impair immune recognition of infected cells (Stanton et al., 2014).

This vast array of genes and mechanisms are summarised in **Table 1.5** and allows HCMV to counteract the host detection and killing mechanisms, allowing it to remain in the host as a lifelong infection, and it is likely that more of these functions will come to light in the future.

1.8 HCMV Gene Families

HCMV, like many DNA viruses, uses many mechanisms in order to generate genetic diversity including host gene capture, gene duplication, substitution, deletion or insertion of nucleotides and large scale genome rearrangements (Prince and Pickett, 2002, Davison et al., 2002). Multiple gene duplications can expand rapidly and have been termed 'accordion expansions', which have been demonstrated in poxviruses under selection pressure (Elde et al., 2012). This process may have resulted in the emergence of a number of distinct multi-gene families in HCMV (Davison et al., 2003a, Chee et al., 1990). The extra gene copies formed by duplications can undergo a divergence in function over time, and some can be subsequently lost. With most families, duplication has resulted in a set of tandem genes, but in others, these duplicate genes have been translocated to other parts of the genome, sometimes with more duplications occurring afterwards, explaining the spread of some gene family members across the genome (Davison, 2011) (**Fig. 1.8**).

Table 1.5: Summary of HCMV immune evasion functions†.

Function category	Genes	References
Modulation of NK ligands	UL16, miR-UL112, UL141*, UL142, US18, US20	(Welte et al., 2003, Cosman et al., 2001, Sutherland et al., 2002, Stern-Ginossar et al., 2007, Tomasec et al., 2005, Tomasec et al., 2000, Fielding et al., 2014, Fielding et al., 2017, Prod'homme et al., 2012, Dunn et al., 2003a)
Other NK evasion strategies	UL83*, UL135, UL18, UL40	(Arnon et al., 2005, Stanton et al., 2014, Prod'homme et al., 2007, Prod'homme et al., 2012, Wang et al., 2002)
Down-regulation of MHC	US2, US3, US6, US11, US8, US10, UL82	(Hsu et al., 2015, Tomazin et al., 1999, Besold et al., 2007, Pande et al., 2005, Noriega et al., 2012b, Lehner et al., 1997, Huber et al., 2002, Tirabassi and Ploegh, 2002, Furman et al., 2002, Trgovcich et al., 2006).
Fc binding proteins	RL11, RL12, RL13, UL119-118	(Cortese et al., 2012, Corrales-Aguilar et al., 2014a, Corrales-Aguilar et al., 2014b)
Cellular homologues	UL111A, US28, UL18, UL33, UL78, UL146, UL22A, UL144	(Kotenko et al., 2000, Chang et al., 2004, Bodaghi et al., 1998, Rosenkilde et al., 2001, Tadagaki et al., 2012, Lüttichau, 2010, Penfold et al., 1999, Scarborough et al., 2017, Wang et al., 2004, Poole et al., 2006, Griffin et al., 2005).
Inhibitors of apoptosis	IE1, IE2, UL36, UL37, UL38, UL141*	(Terhune et al., 2007, Zhu et al., 1995, Goldmacher, 2005, Goldmacher et al., 1999, Seirafian et al., 2014, Skaletskaya et al., 2001, Smith et al., 2013b)
Other antiviral mechanisms	IRS1, TRS1, UL82, UL83*, UL55	(Child et al., 2004, Colberg-Poley, 1996, Ziehr et al., 2016, Browne and Shenk, 2003, Li et al., 2013, Saffert and Kalejta, 2006, Boehme et al., 2004)

† Non-exhaustive list

* Protein has more than 1 role
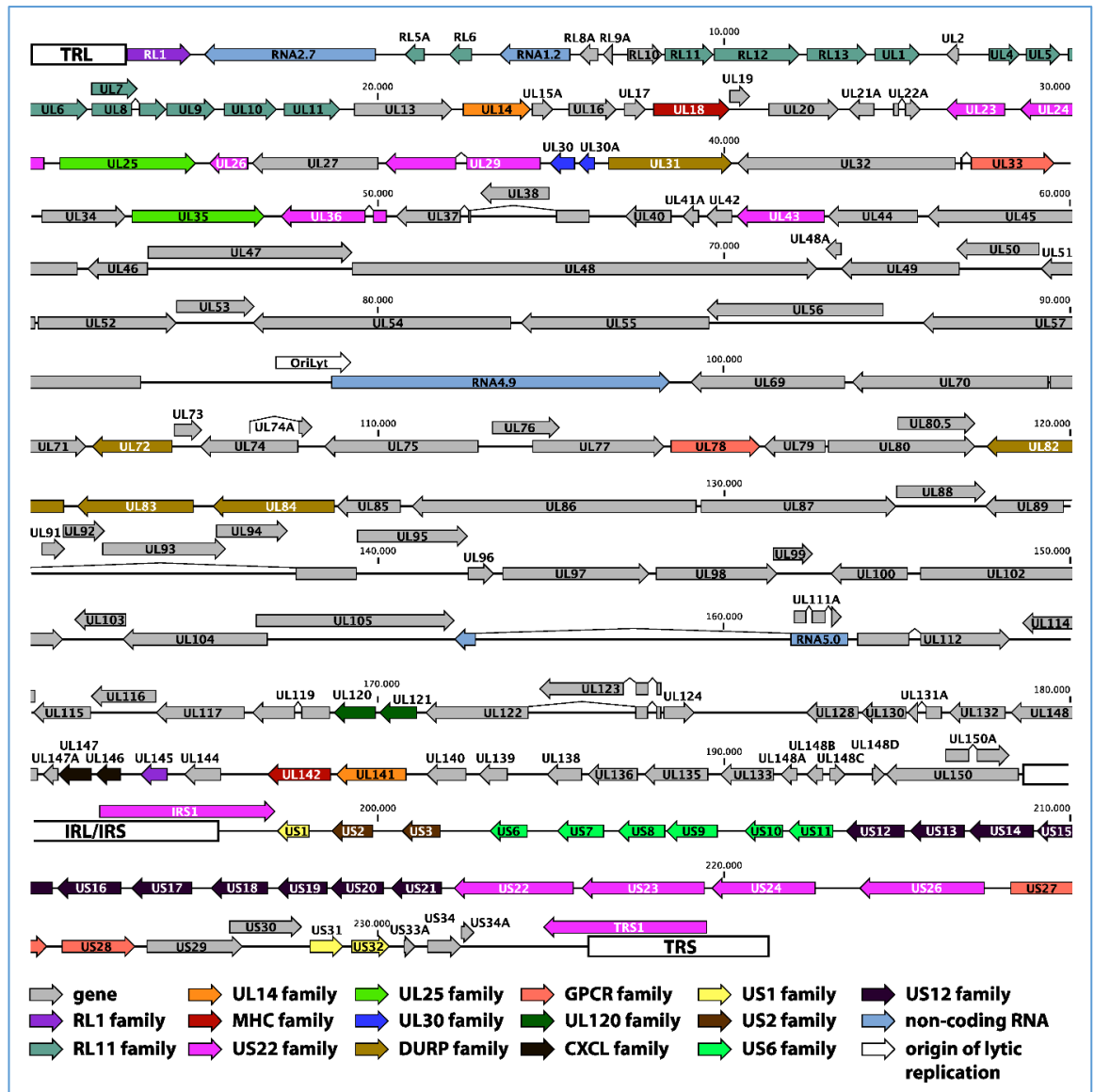
Figure 1.8: The genome of HCMV Merlin with its 15 gene families indicated. Annotation of the Merlin genome including the locations of each gene family. Taken from Steven Sijmons et al. (2014). The single line denotes the dsDNA genome, with nucleotide positions given in basepairs. Genes are represented by block arrows, as well as the 4 non-coding RNAs and the origin of lytic replication. The terminal and internal repeat regions (TRL, IRL/IRS and TRS) are indicated by white boxes. Different gene families are designated with different colour codes, as shown by the key at the bottom of the figure. The UL18, UL82 and UL146 families are also referred to as the MHC, DURP and CXCL families as explained below.

HCMV contains 15 gene families, which is an unusually high number, even compared to other herpesviruses (Davison and Bhella, 2007). The 15 distinct gene families within the HCMV genome include the RL1, RL11, UL14, UL18, UL25, UL30, UL82, UL120, UL146, US1, US2, US6, US12, US22 and G-protein couple receptor (GPCR) families (**Fig. 1.8**) (Sijmons et al., 2014, Gatherer et al., 2011, Chee et al., 1990, Adair et al., 2002, Davison et al., 2003b) . These gene families contain a total of 70 genes, which is over 40% of all canonical genes, and over 55% of non-core genes (Davison, 2011). Nine multigene families were first identified in AD169 (Chee et al., 1990), with additional families being designated later, along with further family members. These additions were added due to advances in sequencing and analytical methods, and the sequencing of additional strains, including low passage clinical strains which retained some genes that heavily passaged strains had lost (Davison et al., 2003b, Dolan et al., 2004, Davison et al., 2003a).

HCMV's gene families range from as few as 2 members (such as the UL14 and UL18 families) to as many as 14 members (the RL11 family). Most of these duplicated and diverged genes have functions that are non-essential for replication *in vitro*, except for UL84 which is essential to the Towne strain for growth in fibroblasts (Dunn et al., 2003b). Some other gene family members do cause growth defects however when deleted in Towne or AD169 (Dunn et al., 2003b, Yu et al., 2003). Some HCMV gene family members are highly variable across different isolates, with the majority of the most variable HCMV genes existing within the RL11 family (Sijmons et al., 2015). As well as gene variability, around 75% of strains tested contained disruptive mutations in 26 genes, including within half of the RL11 family members, with members of the RL1, US6 and US12 families also affected, albeit to a lesser extent (Sijmons et al., 2015). Many gene family members however, are highly conserved in clinical samples, including US1, US12, US22, UL25 and UL82 family members (Sijmons et al., 2015). The UL25, UL82, and US22 gene families are even found within the blocks of core genes that are conserved across beta-herpesviruses (Barry and Chang, 2007). The conservation of gene families and members means they are likely to have important functions, and some members play important roles *in vivo*, and are often involved in immune modulation, including members of the UL14, UL18, US2, US3 and US12 gene families (below). Other gene family members instead have relatively unknown functions. A summary of each gene family can be found below, and in **Table 1.6**.

1.8.1 The RL1 family

The RL1 family is one of the smaller families, consisting of just RL1 and UL145. This is one of the few HCMV gene families in which the genes are not found adjacent in the genome, with

Table 1.6: A summary of HCMV's 15 gene families, including the current members for each family and a brief outline of their functions

Gene family	Family members	Summary of functions*	References
RL1	RL1, UL145	Functions remain uncharacterised	
RL11	RL5A, RL6, RL11, RL12, RL13, UL1, UL4, UL5, UL6, UL7, UL8, UL9, UL10, UL11	Immunomodulatory roles including disruption of T cell activation (UL10, UL11), reduction of downstream signalling pathways, and affecting pro-inflammatory cytokine production (UL7). RL11, RL12 and RL13 are viral Fcγ receptors.	(Bruno et al., 2016, Gabaev et al., 2011, Engel et al., 2011, Corrales-Aguilar et al., 2014b, Lilley et al., 2001, Zischke et al., 2017)
UL14	UL14, UL141	UL141 targets NK activating receptor ligands and TRAIL death receptors to protect against NK cell killing. UL14 has been suggested to impair cell adhesion.	(Tomasec et al., 2005, Prod'homme et al., 2010, Wilkinson et al., 2008, Hsu et al., 2015, Smith et al., 2013b, Cochrane, 2009)
UL18 (or MHC)	UL18, UL142	MHC-I homologues that provide protection from NK cell killing. UL18 achieves this by binding to LIR1 inhibitory receptor, and UL142 possibly down-modulates the NKG2D ligand MICA.	(Chapman et al., 1999, Prod'homme et al., 2007, Wills et al., 2005, Chalupny et al., 2006)
UL25	UL25 (pp85), UL35	UL25 is a tegument protein. UL35 and UL35a appear to interact with pp71 and USP7. UUL35 remodels PML bodies and degrades BclAF1	(Battista et al., 1999, Salsman et al., 2012, Salsman et al., 2011, Lee et al., 2012)
UL30	UL30, UL30A	Functions remain uncharacterised	(Salsman et al., 2008)
UL82 (or DURP)	UL82 (pp71), UL83(pp65), UL84, UL31, UL72	UL82 and UL83 are tegument proteins that affect the cell cycle, modulate gene expression and inhibit antiviral signalling pathways. UL84 is involved in transcriptional activation, DNA replication and viral growth.	(Cantrell and Bresnahan, 2006, Trgovcich et al., 2006, Bresnahan and Shenk, 2000, Arnon et al., 2005, Browne and Shenk, 2003, Fu et al., 2017, Li et al., 2013, Gao et al., 2008, Spector and Yetming, 2010)
UL120	UL120, UL121	Functions remain uncharacterised.	
UL146 (or CXCL)	UL146, UL147	Predicted CXC-chemokines, with UL146 a functional viral homologue of CXCL1 (vCXCL1) that targets CXCR1 and CXCR2. UL147 remains uncharacterised.	(Lüttichau, 2010, Penfold et al., 1999, Scarborough et al., 2017)

Gene family	Family members	Summary of functions*	References
US1	US1, US31, US32	Functions remain uncharacterised	
US2	US2, US3	Both retain or degrade MHC class I and class II proteins to down-regulate their expression on the cell surface. US2 can also reduce integrin signalling, cell adhesion and cell migration through an array of ligands	(Noriega et al., 2012b, Pande et al., 2005, Hsu et al., 2015, Tomazin et al., 1999, Besold et al., 2007, Halenius et al., 2011, Jones and Sun, 1997).
US6	US6, US7, US8, US9, US10, US11	All members regulate MHC molecules in some way, by retaining or degrading them (US9, US11), delaying their exit (US10) or binding them on the surface (US8). US6 inhibits the movement of MHC peptides by binding to the transporter associated with antigen processing (TAP) complex. US7 remains uncharacterised.	(Lehner et al., 1997, Halenius et al., 2011, Besold et al., 2007, Hsu et al., 2015, Furman et al., 2002, Park et al., 2010, Seidel et al., 2015)
US12	US12,US13, US14,US15, US16, US17, US18, US19, US20, US21	The whole family regulate cellular immune ligands, often by targeting them for degradation, and US12, US14, US18 and US20 are NK evasion functions. US16, US18 and US20 are also involved in tropism and US17 can regulate the immune response by altering the content of the virion.	(Cavaletto et al., 2015, Fielding et al., 2017, Luganini et al., 2017, Bronzini et al., 2012, Fielding et al., 2014, Gurczynski et al., 2014, Hai et al., 2006)
US22	US22,US23, US24,US26, UL23, UL24, UL28, UL29, UL36, UL43, TRS1 and IRS1	The majority are tegument proteins, and some have roles in immune evasion; UL36 is a viral inhibitor of caspase-8-induced apoptosis, and TRS1 and IRS1 are antagonists to the antiviral protein kinase R (PKR).	(Adair et al., 2002, Colberg-Poley, 1996, Skaletskaya et al., 2001, Child et al., 2004, Ziehr et al., 2016)
GPCR	UL33, UL78, US27 and US28	All members are chemokine receptor homologues, with US28 able to bind to and sequester a broad range of chemokines. The functions of the remaining family members remain uncharacterised.	(Scarborough et al., 2017, Tadagaki et al., 2012, Bodaghi et al., 1998, Kledal et al., 1998)

* See section of each individual gene family for more details of their functions and other characterisation

RL1 found in the terminal repeat long region, and UL145 found in the unique long region of the genome. Little is known about this gene family and neither of these members have yet to be assigned a function. Only one strain of HCMV (0.8% of strains) studied contained a mutation in RL1 (Sijmons et al., 2015). RL1 is also subjected to higher levels of positive selection than would be expected (Sijmons et al., 2015), so there is a chance that RL1 may interact with host antiviral mechanisms.

1.8.2 The RL11 family

The RL11 family on the other hand is one of the larger families, with 14 members; RL5A, RL6, RL11, RL12, RL13, UL1, UL4, UL5, UL6, UL7, UL8, UL9, UL10 and UL11. They are found near the left terminus of the genome, and members *RL11-UL11* are arranged in tandem except for presence of 2 unrelated genes on the opposing strand (*UL2* and *UL3*) (**Fig. 1.8**) (Davison et al., 2003a). Most members of the RL11 family are predicted to be surface glycoproteins, and all contain a conserved RL11D domain that has homology with human adenovirus E3 membrane glycoproteins CR1 domain (Davison et al., 2003a). Members also generally have a signal peptide, and a transmembrane domain, however not all RL11 proteins contain all 3 of these characteristics; with RL5A and RL6 (found only in HCMV) for example, containing the potential N-linked glycosylation sites but lacking both signal peptides and transmembrane domains (Davison et al., 2003a, Shikhagaie et al., 2012). Hypervariability is a common trait for the family, with RL12, RL13, UL1, UL6, UL9, UL11 and RL6 all shown to have great variability (Dolan et al., 2004, Davison et al., 2003a, Sijmons et al., 2015). As mentioned, there are also an unusually high incidence rate of disrupting mutations found within this family, both *in vivo* and in adaptation to cell culture (Sijmons et al., 2015, Stanton et al., 2010, Dargan et al., 2010, Yu et al., 2002).

Currently 3 members of this gene family are known to function as Fcγ receptors homologues, with gpRL11, gpRL12 and gpRL13 binding to the Fc portion of human IgG which may contribute to HCMV immune evasion (Cortese et al., 2012, Corrales-Aguilar et al., 2014a, Lilley et al., 2001). gpUL7 mediates adhesion to leukocytes and attenuates their production of pro-inflammatory cytokines (Engel et al., 2011). gpUL7 also results in the production of IL-6 by acting as a ligand for the Fms-like tyrosine kinase 3 receptor (Flt-3R) and appears to mediate cellular differentiation of myeloid cells (Crawford et al., 2018). gpUL10 and gpUL11 also have immune modulatory roles, disrupting T cell activation and function, reducing downstream signaling pathways, and affecting pro-inflammatory

cytokine production (Bruno et al., 2016, Gabaev et al., 2011, Zischke et al., 2017). Other proteins, such as gpUL6, currently have no known functions (Sekulin et al., 2007).

1.8.3 The UL14 family

The UL14 family consists of UL14 and UL141, which were defined as a new gene family due to their amino acid homology (Davison et al., 2003b). gpUL141 targets the NK activating receptor ligands CD155 and CD112, providing significant protection against NK cell killing (Tomasec et al., 2005, Prod'homme et al., 2010, Wilkinson et al., 2008). It was found to cooperate with US2 in order to achieve more efficient degradation of specific target cell surface proteins e.g. CD112, and was also shown to target the TRAIL death receptors to protect from TRAIL-dependent NK cell killing (Smith et al., 2013b, Hsu et al., 2015). pUL14 is an uncharacterised homologue but has been suggested to impair cell adhesion (Cochrane, 2009).

1.8.4 The UL18 family

The UL18 family consists of UL18 and UL142, and is known as the MHC family, as each member contains an MHC-I domain, with both members shown to provide protection from NK cell-mediated cytolysis. gpUL18 functions as an MHC class I homolog, binding to the LIR-1 NK inhibitory receptor with >1000 times greater affinity than the endogenous MHC class I ligands (HLA-A, -B, -C, -E, -F and -G), blocking NK-mediated killing (Chapman et al., 1999, Prod'homme et al., 2007). gpUL142 on the other hand, down-modulates the NKG2D ligand MICA (except for the truncated MICA*008 allele) by retaining it in the cis-golgi, also providing a role in NK evasion (Ashiru et al., 2009, Chalupny et al., 2006, Wills et al., 2005). However, in the context of a productive HCMV infection, while gpUL142 provides resistance to NK attack, it does not modulate cell surface expression of MICA (H. Elasifer, personal communication and Fielding et al. (2014)) so the mechanism by which gpUL142 promotes NK evasion is thus uncertain.

1.8.5 The UL25 family

The UL25 family contains just 2 members, UL25 and UL35, with ppUL25 (pp85) encoding a structural tegument protein located in the virion assembly compartment (vAC) (Battista et al., 1999). *UL35* encodes 2 proteins, a full-length 75 kDa protein located in the vAC, and a 22 kDa protein (pUL35A) located in the nucleus (Liu and Biegalke, 2002). Both ppUL25 and the full-length pUL35 protein are packaged into virions and are also found in dense bodies, of which ppUL25 makes up 13% of the total protein amount (Varnum et al., 2004, Liu and Biegalke, 2002, Baldick and Shenk, 1996). *UL25* is dispensable for growth in fibroblasts,

however the *UL35* deletion mutant has a moderate growth defect (Dunn et al., 2003b). Both pUL35 and pUL35a interact with pp71 and it has been suggested that the pUL35A protein may modulate expression of the major IE gene expression by inhibiting activation of the promoter by pp71 (ppUL82) (Liu and Biegalke, 2002). Both pUL35 and pUL35a appear to also interact with each other and the ubiquitin-specific protease USP7 (Salsman et al., 2012, Salsman et al., 2011). pUL35 is able to remodel promyelocytic leukemia (PML) nuclear bodies and is implicated in contributing to viral replication through the manipulation of host responses, including the DNA damage response (Salsman et al., 2012) and degrading the viral restriction factor Bcl-2 associated factor 1 (BclAF1)(Lee et al., 2012).

1.8.6 The UL30 family

The 2 members of this family, UL30 and UL30A, are adjacent in the UL region of the genome, and are conserved among primate cytomegaloviruses (Davison, 2010). The *UL30-UL32* region has multiple overlapping transcripts throughout it, including at least eight mRNAs, and it is predicted that the *UL30* gene is transcribed from the complementary strand (Ma et al., 2013). *UL30A* is related to *UL30*, and is located on the 0.55 kb transcript, with its potential coding region located between the transcriptional initiation site and the translational initiation codon of *UL30*. This *UL30A* region is conserved in Old World primate CMV members but likely uses a non-ATG codon (Davison, 2010, Gatherer et al., 2011). pUL30 was found in sub-organelles of the nucleus, causing the loss or disruption of cajal bodies, which are the main sites for the assembly of small nuclear ribonucleoproteins (RNPs). This may mean that pUL30 plays a role in inhibiting transcription or RNP maturation (Salsman et al., 2008). pUL30 also causes a significant decrease in the number of promyelocytic leukemia (PML) bodies (also known as nuclear domain 10s; ND10s) which are involved in host cell processes such as apoptosis and the DNA damage response, (reviewed by Rabellino and Scaglioni (2013)), disrupting them in a similar fashion to IE1 (Wilkinson et al., 1998). Moreover, PML bodies are known to suppress lytic viral infections (Saffert and Kalejta, 2006, Tavalai et al., 2006) so pUL30 therefore likely reduces the host's suppression on viral infection. Correspondingly, the Towne *UL30* deletion virus has a severe growth defect in fibroblasts (Dunn et al., 2003b)

1.8.7 The UL82 family

The UL82 family consists of UL31, UL72, UL82, UL83, and UL84. They are known as the DURP family as they are deoxyuridine triphosphatase (dUTPase)-related proteins (DURPs), although it is unlikely that they retain this enzymatic activity (Davison, 2013, Davison and Stow, 2005). It is probable that *UL82*, *UL83*, and *UL84* arose from the same ancestral gene

that evolved from host capture of a dUTPase gene, followed by a duplication and subsequent loss of function (Davison and Stow, 2005). ppUL82 (pp71) and ppUL83 (pp65) are both tegument proteins found in high quantities, with pp65 being the most abundant protein in virions (Fu et al., 2017, Varnum et al., 2004). Both proteins are conserved across β-herpesviruses, although pp65 is dispensable for growth *in vitro*, whereas the UL82 deletion mutant has severe growth defects (Dunn et al., 2003b, Fu et al., 2017). Both function as immune evasion proteins, playing a variety of roles from affecting the cell cycle (Cantrell and Bresnahan, 2006, Trgovcich et al., 2006), modulating gene expression (Bresnahan and Shenk, 2000) and inhibiting antiviral signalling pathways (Noriega et al., 2012b, Halenius et al., 2011, Fu et al., 2017, Arnon et al., 2005, Browne and Shenk, 2003, Li et al., 2013). pUL84 has been shown to interact with ppUL83 and is also present in virions, but not in large quantities (Varnum et al., 2004, Gao et al., 2008). pUL84 shuttles from the nucleus to the cytoplasm, is involved in transcriptional activation, DNA replication and viral growth, and is essential for growth of both Towne and AD169 strains, although not TB40/E (Dunn et al., 2003b, Gao et al., 2008, Spector and Yetming, 2010, Davison and Stow, 2005).

UL31 and UL72 members belong in a separate group, with little known about either member except that pUL31 localizes to the nucleolus in uninfected cells (Salsman et al., 2008). Both members cause moderate growth defects in fibroblasts when deleted, despite not being essential for growth (Dunn et al., 2003b). *UL72* comprises a core HCMV gene and has been revealed not to be an activate dUTPase despite being an ortholog of functional dUTPases (Davison and Stow, 2005, Caposio et al., 2004).

1.8.8 The UL120 family

The UL120 family consists of just UL120 and UL121, found adjacent in the major IE (MIE) region of the genome, alongside UL123 (IE1) (Grey et al., 2007). They are both putative membrane glycoproteins that are distantly related to each other, and comprise one of the newer gene families (Dolan et al., 2004). It is possible that they arose via duplication of the *UL119* ancestor gene as there is marginal conservation between them in a portion of the immunoglobulin domain, making UL119 another possible family member (Davison and Bhella, 2007). UL120 and UL121 constitute 2 of the target sequences for the viral microRNA miR-UL112-1. Their function is currently unknown, but miR-UL112-1 regulates the expression of genes involved in viral replication, and their location in the MIE has suggested

that *UL120* and *UL121* may encode exons within the MIE family of transcripts (Grey et al., 2007).

1.8.9 The UL146 family

The UL146 family consists of UL146 and UL147 and are known as the CXCL family as the proteins are either known or predicted CXC-chemokines (Penfold et al., 1999, Dolan et al., 2004). CXC receptors 1 (CXCR1 and CXCR2) are the main host chemokines receptors expressed on neutrophils and they play a major role in the inflammatory response. UL146 is a viral homologue of the endogenous CXC ligand CXCL1 (vCXCL1), which targets CXCR1 as a selective agonist, and also CXCR2 with lower affinity and potency (Lüttichau, 2010, Penfold et al., 1999). gpUL146 has known chemokine function and HCMV may use gpUL146 to attract neutrophils to infected cells after which they could act as carriers of the virus to uninfected locations (Lüttichau, 2010, Penfold et al., 1999). *UL146* is one of the most hypervariable HCMV genes and is highly divergent throughout its length, and exists as 14 variants or alleles, which may be driven by differing gene function (Davison, 2011, Stanton et al., 2005). pUL147 (designated vCXCL2) has yet to be fully characterised but has less homology to CXC chemokines and is less likely to have chemotactic ability (Scarborough et al., 2017, Penfold et al., 1999).

1.8.10 The US1 family

The US1 family includes the members US1, US31, US32 (Rigoutsos et al., 2003). Little is known about their function but all 3 members are dispensable for growth *in vitro* (Dunn et al., 2003b). This gene family are found in HCMV, RhCMV and CCMV amongst others, and each members contains two copies of a motif in their N-terminal regions which may coordinate a metal ion (Davison and Bhella, 2007). pUS32 is a late protein that co-localises with PML or ND10s and its expression results in altered size and shapes of nuclear bodies, although the mechanism and reason are unknown (Salsman et al., 2008, Strang, 2015).

1.8.11 The US2 family

The US2 family members, US2 and US3, are 50% similar at the amino acid level and both glycoproteins are expressed early in infection (Jones and Sun, 1997). As mentioned in **Section 1.7.4.1**, gpUS2 and gpUS3 both down-regulate both MHC class I and class II expression on the cell surface (Noriega et al., 2012b, Pande et al., 2005, Hsu et al., 2015, Tomazin et al., 1999, Besold et al., 2007). gpUS3 functions by retaining MHC class I heavy chains within the ER and can also interfere with the chaperone tapasin that controls peptide loading (Halenius et al., 2011, Noriega et al., 2012b). gpUS2 on the other hand

causes the rapid destabilization and degradation of the retained MHC proteins but can also reduce integrin signalling, cell adhesion and cell migration through an array of ligands (Hsu et al., 2015, Tomazin et al., 1999, Jones and Sun, 1997).

1.8.12 The US6 family

The US6 family currently consists of 6 predicted glycoproteins US6 to US11, which cluster separately as a family but have partial amino acid homology with the US2 family (Huber et al., 2002, Pande et al., 2005). Multiple transcripts are produced across the family region with members sometimes present on more than 1 transcript (Jones and Muzithras, 1991). The deletion of both US2 and US6 families (US2-11) was able to completely prevent the presentation of IE1 antigens (Besold et al., 2007). Most US6 family members work to down-regulate MHC I expression to avoid immune detection or are involved in antigen presentation in some way (Huber et al., 2002, Noriega et al., 2012a). gpUS6 inhibits the movement of peptides across the lumen of the ER membrane by binding to the transporter-associated with antigen processing (TAP) complex, and US11 retains and degrades MHC molecules (Huber et al., 2002, Lehner et al., 1997, Halenius et al., 2011, Besold et al., 2007, Hsu et al., 2015). pUS8 is able to bind to MHC I on the surface and pUS10 can delay the exit of MHC I and down-regulating the cell surface expression of HLA-G (Furman et al., 2002, Huber et al., 2002, Park et al., 2010). pUS9 downregulates the glycosylphosphatidylinositol (GPI)-anchored MICA*008 allele by targeting it for proteasomal degradation, although it does not appear to affect the overall MICA*008 levels during HCMV infection (Seidel et al., 2015). pUS7 is also predicted to have an immune evasion function but remains largely uncharacterised. The US6 family are not required for growth of HCMV in fibroblasts *in vitro,* and mutations that naturally occur in *US6*, *US7*, and *US9* point to functional redundancy of the family (Sijmons et al., 2015, Dunn et al., 2003b).

1.8.13 The US22 family

The US22 family is the second largest HCMV gene family, as well as the most highly conserved, also being encoded by other β-herpesviruses (Lesniewski et al., 2006). It contains the members US22, US23, US24, US26, UL23, UL24, UL28, UL29, UL36, UL43, TRS1 and IRS1, with at least 7 members documented as tegument proteins, indicating that the rest may play a similar role (Adair et al., 2002). Accordingly, most members are found in HCMV virions and some can also be found in aggregates resembling dense bodies (Adair et al., 2002).

Their conserved US22 sequence motifs may be tegument signals that have evolved to allow for more diverse functions, and as such, they play a role in a variety of processes, including immune evasion (Lesniewski et al., 2006). pUL36 encodes a cell death suppressor and is a viral inhibitor of caspase-8-induced apoptosis (vICA) by binding to caspase-8 and preventing its activation and thereby Fas-mediated apoptosis (Skaletskaya et al., 2001). pTRS1 and pIRS1 are highly homologous proteins that are also involved in the evasion of the hosts antiviral responses (Ziehr et al., 2016, Colberg-Poley, 1996, Child et al., 2004), both proteins designated as antagonists to the antiviral protein kinase R (PKR) which normally limits viral protein translation and synthesis upon binding to viral dsRNA. Both are likely to be transcriptional transactivators and may cooperate with other IE proteins, and expression of either is sufficient to bind and inactivate PKR, thereby allowing for efficient synthesis and replication to take place (Ziehr et al., 2016, Child et al., 2004). The deletion of either is most likely not detrimental due to their common functions or targets, however the deletion of both results in a replication-deficient virus (Ziehr et al., 2016, Jones and Muzithras, 1991, Strang, 2015). Many US22 family members still have unknown function, and although all are non-essential for growth in AD169 or Towne strains *in vitro*, deletion of *US24* does cause a growth defect, with a delay and a decrease in the expression of IE, E and L proteins (Feng et al., 2006, Dunn et al., 2003b).

1.8.14 The GPCR family

The G-protein coupled receptor (GPCR) family includes the members UL33, UL78, US27 and US28, 2 of which are found in the UL region and 2 in the US region of the genome**.** These 4 proteins are predicted to encode 7-transmembrane proteins which contain the key hallmark features of chemokine receptor homologues (Scarborough et al., 2017). US27 and US28 exhibit homology to, and likely arose from, a common ancestor with human chemokine receptor CX3CR1 (Kledal et al., 1998, Scarborough et al., 2017). pUS28 binds to, and can sequester, a broad range of chemokines including RANTES, MIP-1α, and MCP-1, although it is much more specific for the CX3CR1 ligand fractalkine (Kledal et al., 1998, Bodaghi et al., 1998). It can subsequently activate the major MAP-kinase pathways, as well as activating the transcription of NF-κB (Rosenkilde et al., 2001) and is one of the few genes expressed during "latency" (Scarborough et al., 2017). pUS27 is a putative chemokine receptor, and although it has no known ligands to date, it appears to directly enhance the calcium signalling activity of a human chemokine receptor, CXCR4 (Arnolds et al., 2013). pUL33 and pUL78 can also regulate cellular chemokine receptors CXCR4 and CCR5 through receptor heteromerization and are suggested to be important for virulence, however their

functions and mechanisms are also unknown (Rosenkilde et al., 2001, Tadagaki et al., 2012). The importance of US27 and US28 is indicated through their high conservation of across different HCMV strains, with a minimum of 94.74% and 97.46% identity respectively (Scarborough et al., 2017). As homologues of chemokine receptors, the GPCR family could have a range of functions including protecting the receptor from recognition by the host immune system, a way to eliminate chemoattractants from the surrounding area, and potential involvement with HCMV cell entry, tissue targeting or attracting specific immune cells for dissemination purposes (Kledal et al., 1998, Bodaghi et al., 1998, Rosenkilde et al., 2001).

1.8.15 The US12 family

The US12 family is comprised of 10 tandemly aligned members US12 to US21, arranged across 9 kb in HCMV's unique short (US) region (Dolan et al., 2004, Chee et al., 1990). The US12 family is one of the larger multi-gene families, and its members are produced across 3 transcriptional cassettes, US21, US20-US18 and US17-US12 (Guo, 1993, Lu et al., 2016). The US12 family are highly conserved across HCMV strains and in the CMVs of higher primates, with the ancestral virus likely to have encoded at least 8 of the US12 family members prior to the divergence of the rhesus and hominoid lineages (Lesniewski et al., 2006). Some primate CMVs are missing some of the US12 family members and others have extra members, with eleven members found in the rhesus macaque CMV (RhCMV) genome, which lacks *US15* and *US16*, but has 4 homologues of *US14* (Hansen et al., 2003). The peripheral members of the group (*US12*, *US13* and *US17-US21*) are the highest conserved members, with *US21* having the most similarity across higher primate CMV strains. *US21* also shows significant overall sequence similarity to three human proteins (lifeguard, CGI-119, and PP1201), so proto-*US21* is suggested to represent the origin of the family (Lesniewski et al., 2006, Holzerlandt et al., 2002). It is thought that the duplication and divergence of proto-*US21* lead to the creation of *US20* which was then important for the development of the rest of the family (Lesniewski et al., 2006).

The US12 family have some homology to G-protein coupled receptor (GPCR) proteins, with most members predicted to have 7 transmembrane (7TM) domains, 6 members having DRY motif related sequences and 5 members containing a motif closely related to a sequence found in the conserved GPCR Frizzled (Lesniewski et al., 2006). Overall, US12 and US14 are the most GPCR-like as they contain the greatest number of similar motifs, whereas US19 is the least similar, containing just 1 GPCR motif (Rigoutsos et al., 2003, Lesniewski et al., 2006). There is no evidence however of the US12 family having any GPCR-

like signalling, and they represent a distinct branch of the GPCR group so have the potential to encode vastly different functions (Lesniewski et al., 2006). There is also shared similarity in a motif between the US12 family and the transmembrane BAX Inhibitor-1 Motif-containing (TMBIM) protein family that includes Bax inhibitor-1 (BI-1) and related 7TM proteins (Lesniewski et al., 2006, Rojas-Rivera and Hetz, 2015). BI-1 inhibits Bax-mediated apoptosis, but to date no apoptosis-related functions have been discovered within the US12 family.

All US12 family members are non-essential for growth in fibroblasts *in vitro,* however US13 does cause a slight growth defect when deleted in fibroblasts (Dunn et al., 2003b), and other members have growth defects when grown in different cell types (**Section 1.8.15.2**). Despite being non-essential, the US12 family are still highly conserved in clinical isolates, with *US13* and *US18* in the top 25 most conserved genes within HCMV (Sijmons et al., 2015, Dunn et al., 2003b, Yu et al., 2003). Most US12 family members also showed no ORF-disrupting mutations within 124 clinical isolates tested, with just 1 strain presenting with a 2 nucleotide (nt) insertion in *US12* and 3 strains demonstrating a 24 nt deletion within *US13* (Sijmons et al., 2015). Their conservation implies that they play important roles *in vivo*, and so far the family have been shown to have a variety of roles in immune evasion, replication, and tropism *in vitro* (**Section 1.8.15.2 and 1.8.15.4**).

1.8.15.1 The localisation of the US12 family

Many US12 family members are expressed early during infection, including pUS19 and gpUS20 (Chambers et al., 1999, Guo, 1993, Cavaletto et al., 2015), with some members accumulating later during infection such as pUS16, pUS17 and pUS18 (Bronzini et al., 2012, Das et al., 2006). Proteins pUS14, pUS16, pUS17 and pUS18 from the TR or AD169 strains displayed localisation with the virion assembly compartment (vAC) at some point during infection, mostly accumulating there at later time-points (Das and Pellett, 2007, Das et al., 2006, Bronzini et al., 2012). AD169 pUS14 and pUS18 also had cytoplasmic distributions, with pUS14 presenting in a granular manner, and pUS17 also having nuclear localisation of its C-terminal segment (Das et al., 2006, Das and Pellett, 2007, Bronzini et al., 2012). TR pUS20 however, was shown to localise to ER-derived membranes (Cavaletto et al., 2015). Although it is obvious that all US12 family members do not share the same localisation, there does appear to be a trend towards an association with the vAC, although these locations may differ in other strains and in other cell types (Das and Pellett, 2007, Das et al., 2006, Bronzini et al., 2012).

1.8.15.2 The role of the US12 family in replication and tropism

Currently, three members have been shown to play roles in HCMV replication and are involved with the ability of HCMV to grow in different cell types, with some deletion mutants failing to express immediate-early (IE) genes upon infection, abrogating growth of the virus (Bronzini et al., 2012, Hai et al., 2006, Dunn et al., 2003b). Deletion of the AD169 *US16* gene for example, was found to cause a major growth defect in both endothelial cells and epithelial cells, whilst replication was normal in fibroblasts (Bronzini et al., 2012). US16 appears to function during the final stages of virus maturation, reducing the virion content of the pentamer complex that is required for efficient entry into endothelial and epithelial cells, thereby reducing viral growth in those cells types only (Luganini et al., 2017, Bronzini et al., 2012). Another major growth defect was observed with the deletion of *US18* from the Towne strain, this time observed in cultured human gingival tissues derived from human keratinocytes (Hai et al., 2006). There was reduced viral growth in this cell type compared to fibroblasts, with the mutant appearing to be deficient in infecting the tissue as well as replicating within them, blocked at a stage prior to (or at) IE gene expression (Hai et al., 2006). *US20* was instead required for efficient growth in endothelial cells at a stage after attachment and entry, but the *US20* deletion mutant replicated normally in fibroblasts and epithelial cells (Cavaletto et al., 2015). Other differential effects have yet to be discovered, but the US12 family themselves are also differentially regulated in different cell types, with at least 5 members (*US12, US14, US18, US19 and US20*) having differential genes expression between astrocytoma cells, fibroblasts and retinal pigmented epithelial cells (Towler et al., 2012).

1.8.15.3 The role of the US12 family in regulating the cellular and virion proteome

HCMV modifies the host's cellular gene expression in order to evade the host's defences, with the US12 family targeting multiple host plasma membrane proteins (Weekes et al., 2014, Fielding et al., 2017). The US12 family caused major changes in the expression levels of over 80 proteins, some of them targeted for lysosomal degradation by the family (Fielding et al., 2017). Pathways significantly affected by the US12 family, included cell adhesion molecules, cytokine-cytokine receptor interactions and natural killer (NK) cell-mediated cytotoxicity pathways (Fielding et al., 2017). Multiple US12 family members have the ability to target more than 1 protein each, with gpUS20 able to regulate 54 cellular targets. Different family members can also target the same proteins; with 29% of plasma membrane proteins regulated >3 fold by 2 or more family members, and 6% by 3 or more family members (Fielding et al., 2017). gpUS20 for example, shares at least 14 of its targets

with pUS18, and the 2 have demonstrated to act in concert (**Section 1.8.15.4**) (Fielding et al., 2017, Fielding et al., 2014).

As well as regulating their target proteins, the US12 family can also affect the expression levels of each other, for example, the deletion of US13 or US15 increases the expression of pUS14, and deletion of US21 increases pUS20 levels (Fielding et al., 2017). This may suggest a compensation mechanism for members that regulate the same targets, or may imply that these proteins form complexes with each other in order to accomplish their roles (Fielding et al., 2014).

The US12 family can also exert effects on the virion proteome. pUS17 not only plays a role in the final stages of virion assembly and egress, but also alters the levels of virion proteins, such as the envelope glycoprotein H (gH) (Gurczynski et al., 2014). pUS17 additionally alters the ratio of infectious to non-infectious particles, with ΔUS17 producing more genome-containing non-infectious particles than its parental virus, whilst producing the same number of infectious virions (Gurczynski et al., 2014). The US16 deletion mutant instead reduces the levels of the pentamer complex on the virion, abrogating entry into epithelial and endothelial cells (Bronzini et al., 2012, Luganini et al., 2017). Other links between US12 family members are virion composition have yet to be established.

1.8.15.4 The role of the US12 family in immune evasion

The modulation of the cellular proteome by HCMV includes the targeting of many cellular immune ligands for degradation, and this may allow HCMV to favourably alter the host immune response. The US12 family's regulation of the NK cell-mediated cytotoxicity pathway for example, includes the ability to downregulate NK activating receptor ligands such as MICA, MICB, ULBP2 and B7-H6 (Fielding et al., 2017, Fielding et al., 2014). Further to this, US12, US14, US18, US20 and US21 have all been identified as NK evasins (Fielding et al., 2017, Fielding et al., 2014). The family either retain their targets intracellularly, or they target them for degradation. pUS18 and gpUS20 for example, both work in concert to target the NKG2D ligand MICA for lysosomal degradation and to target the NKp30 ligand B7-H6 for proteolysis, reducing activation of NK cells in response to the infected cell (Fielding et al., 2017, Fielding et al., 2014).

The US12 family can also influence immune evasion indirectly, with the deletion of US17 causing an alteration in virion composition. The deletion of US17 leads to decreased levels of gH and increased amounts of pp65 being delivered to newly infected cells, causing significantly differential expression of innate and intrinsic immune response-related gene

transcripts. This alters the immune response to HCMV in newly infected cells and differentially regulates the endoplasmic reticulum stress response at 96 hpi (Gurczynski et al., 2014). The fact that some members, especially *US14* and *US18*, seem to be subjected to higher levels of positive selection than would be expected from their diversity supports the idea that they may interact with host antiviral mechanisms (Sijmons et al., 2015).

1.8.15.5 Summary

In summary the US12 family have a variety of roles including tropism (US16, US18 and US20), immune evasion (US12, US14, US18, US20 and US21) and virion composition and maturation (US16, US17) as summarised in **Table 1.7** (Bronzini et al., 2012, Hai et al., 2006, Cavaletto et al., 2015, Luganini et al., 2017, Gurczynski et al., 2014). The US12 family affect the levels of a wide range of host proteins and immune ligands, as well as each other (Fielding et al., 2014, Fielding et al., 2017). Although the localisation and expression levels of certain US12 family members have been determined, these have generally yet to be tested in a clinically relevant virus such as Merlin, with US12, US13, US15, US19 and US21 having little or no functional or characterisation data available.

1.8.16 Aims and hypothesis

The US12 family is one of the largest of the 15 multigene families in HCMV's genome and some US12 family members have been shown to have important roles in HCMV infection in different strains and different cell types (Bronzini et al., 2012, Hai et al., 2006, Cavaletto et al., 2015, Fielding et al., 2014). It is likely that other family members are likely to have similar roles and be similarly important for HCMV. Despite some recent functional studies on the target proteins of the US12 family (Fielding et al., 2017), little is known about the fundamental aspects of the proteins themselves. The aim of this project is therefore to further characterise the US12 family in the clinically relevant strain Merlin, and to give a greater understanding of the basic fundamental aspects of the entire gene family in productive HCMV infection. The intention of the project is also to assess whether these findings give insights into their mechanism of action, including their role in immune evasion, and how the US12 family link into the overall HCMV infection.

The US12 family members appear to play a role in immune modulation by targeting cellular immune receptors for lysosomal degradation (Fielding et al., 2017), and we hypothesise that they may achieve this by trafficking the proteins to the lysosomes themselves. If this is the case, the US12 family members may also be degraded themselves, and this can be

Table 1.7: Summary of the current published knowledge of the US12 gene family members, pertaining to their localisation, tropism and immune evasion functions.

Protein	Localisation	Tropism	Immune evasion
US12			NK evasion function. Regulates cellular immune ligands, including ULBP2 [**a**]
US13			Regulates cellular immune ligands, including MICB [**a**]
US14	vAC and dispersive cytoplasmic (AD169) [**b, c**]		NK evasion function. Regulates cellular immune ligands, including JAM3 [**a**]
US15			Regulates cellular immune ligands, including IL6ST [**a**]
US16	vAC (TR strain) [**d**]	Required for efficient infection of endothelial and epithelial cells [**d,e**]	
US17	N-terminus to the periphery of the vAC and C-terminus to the nucleus and cytoplasm (AD169) [**b, c**]		Affects immune responses through altered virion composition e.g. of pp65 [**f**].
US18	Cytoplasmic, moving to the vAC later on (AD169) [**b, c**]	Required for efficient growth in human gingival tissue [**g**]	NK evasion function. Regulates cellular immune ligands including B7-H6 and MICA in concert with US20 [**a, h**]
US19			
US20	Sub-cytoplasmic ER localisation (TR strain) [**i**]	Required for efficient growth in endothelial cells [**i**]	NK evasion function. Regulates cellular immune ligands, including MICA and B7-H6 [**a, h**]
US21			Possible NK evasion function [**a**].

[**a**]= Fielding et al. (2017); [**b**]= Das and Pellett (2007); [**c**]= Das et al. (2006); [**d**]= Bronzini et al. (2012); [**e**]= Luganini et al. (2017); [**f**]= Gurczynski et al. (2014); [**g**]= Hai et al. (2006); [**h**]= Fielding et al. (2014); and [**i**]= Cavaletto et al. (2015).

tested by observing the effects of adding a lysosomal inhibitor. We hypothesize that some US12 family members may work together in a complex, supported by the fact that multiple proteins can target the same cellular proteins, and identifying their localisations could postulate whether multiple members are found in the same locations. We additionally theorize that any US12 family members that are found in the vAC are also likely to be incorporated into the virion.

In order to achieve these goals, HCMV bacterial artificial chromosomes (BACs) were constructed with a V5 tag added to each US12-US21 gene individually. These viruses were then used to characterise the US12 family members for their expression levels, their intracellular trafficking and their post-translational modifications. This investigation of the US12 family could advance our understanding of HCMV pathogenesis and immune modulation and may potentially allow for the family to be targeted in therapy in the future.

2 Materials and Methods

2.1 Molecular Biology

2.1.1 Growth of *E. coli* SW102 cultures

The HCMV Merlin bacterial artificial chromosome (BAC) was propagated using *Escherichia coli* SW102 bacteria for fast growth and easy BAC manipulation. *E. coli* SW102 were utilised as they contain lambda phage *Red* recombination genes under temperature sensitive expression. These genes could be induced by incubating the bacteria at 42°C allowing homologous recombination of DNA to take place, and the bacteria otherwise grown at 32°C to avoid inappropriate recombination.

2.1.2 Selection media and plates

E. coli SW102 containing the Merlin BAC were grown overnight in 5ml Luria-Bertani (LB) broth at 32°C overnight in a shaking incubator (Stuart orbital incubator) as standard. All media was autoclaved for sterility and allowed to cool to below 50°C before antibiotics were added. Chloramphenicol was used as standard to select for SW102 bacteria that contained the HCMV BAC. Transformed bacteria were grown on LB agar plates when selecting for single colonies or correctly recombineered BACs. Different selection cassettes (**Section 2.1.4.1**) required different antibiotics and chemicals which were prepared as described in **Table 2.1**. The LB agar was then poured into sterile petri dishes (20 ml each, Fisher, PDS-140-050F) under the sterility of a Bunsen burner and allowed to solidify before being stored upside down at 4°C.

GalK recombineering methods required M63 minimal media plates, and 2-deoxy-galactose (DOG) was utilised as the carbon source as only the second round of recombineering was undertaken (**Section 2.1.4.1**).

LB broth media: 10g of LB powder (Melford Biolaboratories Ltd, L1703) dissolved in 500ml dH$_2$0

LB agar: add 7.5g agar (Sigma-Aldrich) to 500ml LB broth

Sucrose LB agar plates: 5g tryptone (Thermo Fisher), 2.5g yeast extract (Sigma-Aldrich), 7.5g agar and 25g sucrose (Fisher, BP2201) dissolved in 500 ml dH$_2$0.

M63 minimal media plates: 15 g agar dissolved in 800 ml dH$_2$O and autoclaved. Once cooled a little, add 200 ml 5X M63 medium and 1 ml 1 M Magnesium sulfate heptahydrate (MgSO4·7H2O). Top up to 1l with dH20 if required. Once cooled to 50°C, add 10 ml 2-

Table 2.1: Supplementary antibiotics and chemicals added to LB broth and agar plates

Antibiotic/chemical	Main stock:	Working stock:	Use at:
Chloramphenicol	Powder (Doehringer)	Made to 12.5 mg/ml by dissolving 625 mg in 50 ml ethanol	1:1000 (12.5 µg/ml final concentration)
Ampicillin sodium	Powder (Duchefa Biochemie)	Made to 100mg/ml by dissolving 1g ampicillin sodium sulphate in 10 ml dH$_2$0, then filter sterilized*	1:2000 (50µg/ml final concentration)
Kanamycin monosulphate	Powder (Melford)	Made to 15 mg/ml by dissolving 150 mg in 10 ml dH$_2$0, then filter sterilized*	1:1000 (15µg/ml final concentration)
Streptomycin sulphate	Powder (Melford	Made to 200mg/ml by dissolving 2g of Streptomycin sulphate in 10 ml dH$_2$0, then filter sterilized*	1:500 (400µg/ml final concentration)
5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-gal)†	Powder (Melford)	Made to 40mg/ml by dissolving 400mg of X-gal in 10ml 100% DMSO, and stored in foil	1:500 (80 ug/ml final concentration)
Isopropyl β-D-1-thiogalactopyranoside (IPTG)†	Powder (Melford)	Made to 100mM by dissolving 0.238g IPTG in 10ml dH$_2$0, then filter sterilized*	1:500 (200 nM final concentration)

*Once dissolved, the solution must be filter sterilised by passing through a 0.22µm filter

† Chemicals used to aid the blue/white screening of recombinants (**Section 3.1, Figure 3.3**).

deoxy-galactose (DOG) (0.2%), 5 ml biotin (1 mg), 4.5 ml leucine (45 mg), and 500 ml chloramphenicol (12.5 mg/ml).

5X M63: 10 g of ammonium sulfate ((NH4)2SO4), 68 g potassium dihydrogen phosphate (KH2PO4) and 2.5 mg ferrous sulfate heptahydrate (FeSO4·7H2O) in 1l of dH_2O. Adjust to pH7 with potassium hydroxide (KOH).

2.1.3 Generation of glycerol stocks

Long term storage stocks of each clone were made from overnight cultures of SW102 containing the required HCMV Merlin BAC, with 500 ml of LB culture added to 75 µl glycerol (15%) in 1 ml screw cap tubes with seals (VWR, 16466-054). Once mixed, the stocks could be frozen at -70^oC for future use, providing a reproducible source of BAC. To propagate from these, an inoculation loop was used to take a scrape of the frozen glycerol stock which was added to 5 ml of LB with appropriate selection antibiotics and grown overnight at 32^oC in a shaking incubator.

2.1.4 Recombineering

Recombination-mediated genetic engineering or 'recombineering' is a method of altering the virus BAC genome utilising the lambda red genes of E. coli SW102. To construct US12 family V5-tagged virus BACs, 2 rounds of homologous recombination were required **(Section 3.1, Figure 3.2)**. Briefly, the first round of recombineering involved inserting a selection cassette after the gene of interest (US12 family member), by adding in a PCR product of the cassette that contains homology arms adjacent to the gene. In the second round of recombineering, the cassette was then swapped for oligos of the V5 tag that is flanked by homology arms of the gene of interest and downstream homology.

2.1.4.1 Selection cassettes

A variety of selection cassettes were used for recombineering, the SacB, Rpsl and GalK cassettes. Each selection cassette contains an antibiotic resistance gene which allows for positive selection of clones that have successfully inserted the cassette during the first round of recombineering, along with a gene used for negative selection of clones that have subsequently swapped out the cassette for the V5 tag during the second round of recombineering (**Table 2.2**). The Rpsl cassette (*KanR/SmS/LacZ/Rpsl*) encoded for kanamycin resistance for positive selection and streptomycin sensitivity for negative selection, which inhibited growth for clones that had failed to remove the cassette. The

Table 2.2: Selection cassettes utilised for the production of V5-tagged genes in the HCMV merlin BAC by recombineering

Cassette	Resistance/ positive selection genes	Sensitivity/ negative selection genes	Other genes	Reference
Rpsl (*KanR/SmS/ LacZ/Rpsl*)	Kanamycin resistance marker (*neo^R/KanR*)	Streptomycin sensitivity (*Rpsl+*)	LacZα for white/blue screening (*lacZ*)	(Sung et al., 2001)
SacB (*AmpR/LacZ/ SacB*)	Ampicillin resistance gene and lacZ (*amp^R*)	SacB for sucrose sensitivity (*SacB*)	LacZα for blue/white screening (*lacZ*)	(Stanton et al., 2008)
GalK	Ability to use galactose as the only carbon source (g*alE, galT, galK* and *galM*)	2-deoxy-galactose (DOG) sensitivity/toxic build up (galK)	-	(Warming et al., 2005)

SacB cassette (*AmpR/LacZ/SacB*) encoded for ampicillin resistance for positive selection and levansucrase for negative selection, inhibiting growth for cassette-containing clones in the presence of sucrose. Both Rpsl and SacB cassettes also contained the *LacZ* gene which encoded for β-galactosidase (**Table 2.2**). Isopropyl β-D-1-thiogalactopyranoside (IPTG) in the selection media stimulated lacZ transcription, and the induced β-galactosidase cleaved its chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-*d*-galactopyranoside (X-gal). X-gal is a colourless analog of lactose, and it's cleavage resulted in the formation of a blue insoluble pigment (5,5'-dibromo-4,4'-dichloro-indigo), and allowed for the identification of successfully transformed colonies by colour, termed 'blue/white screening' (**Section 3.1, Figure 3.3**).

The GalK cassette encoded for galactokinase (galK) which phosphorylates galactose and its derivatives and was used for both positive and negative selection. The GalK cassette allowed SW102 bacteria (which contain a deletion of galK) to grow on minimal media with galatose as the only carbon source, allowing for positive selection. By exchanging the galactose in the media for its analog 2-deoxy-galactose (DOG), the presence of galK would lead to a toxic build-up of the product 2-deoxy-galactose-1-phosphate, allowing for negative selection (**Table 2.1**). Appropriate antibiotics for each cassette were made up and added to selection media or agar plates (**Section 2.1.2**).

2.1.4.2 Preparation of competent bacteria for recombineering

The first round of recombineering required SW102 that contained the HCMV Merlin pAL1111 BAC, to be grown overnight at 32°C in LB with chloramphenicol. The 2nd round required SW102 that contained the BAC in which the cassette had been inserted at the end of the US12 family gene (e.g. Merlin BAC with US12-Rpsl) which was grown overnight in LB with chloramphenicol and the appropriate selection cassette antibiotic (e.g. kanamycin for Rpsl). 0.5 ml of this overnight culture was added to 25 ml LB with appropriate antibiotics in a falcon tube and was left in a shaking incubator at 32°C until it reached an optical density (OD) of 0.6 at 600nm. A UV spectrophotomer (Pharmacia Ultraspec 3000) was used to measure the OD, and OD 0.6 signalled the exponential phase of bacterial growth. Once this OD was achieved, the falcon tube was heated to 42°C in a waterbath for 15 min, whilst inverting often, to induce the lambda phage genes. The bacteria were then cooled by rocking on ice for 15 min, and kept at ~0°C for all future steps to keep the cells competent in preparation for recombination. The samples were centrifuged at 4000 rpm for 5 min at

0°C, the supernatants discarded, and the pellets re-suspended in 1ml sterile ice-cold water by gentle shaking. These were topped up to 25 ml with cold water and re-centrifuged at 4000 rpm for 5 min. This wash step was repeated once more and after the final centrifugation, the supernatant was discarded and the pellet re-suspended in the small volume of water left in the falcon tube.

2.1.4.3 Transformation of competent bacteria

A 25 µl aliquot of the competent SW102 resuspension was added into a pre-cooled 0.2 cm electroporation cuvette (GeneFlow, E6-0060) with the addition of the DNA of interest- either 4 µl of purified cassette PCR product or 1 µl of the V5 tag oligonucleotides and left on ice for 5 min. The sample was then electroporated at 2.5 kV on program EC2 on a Micropulser (Bio-Rad). For recovery, 1ml LB was added and samples kept at 32°C for 1 hour in the shaking incubator, with a 4 hour recovery in 5ml LB for the SacB negative selection step.

Multiple dilutions of recovered transformed SW102 (generally 20 and 100 µl) were added onto the required LB selection plates (**Section 2.1.4.4**) and spread using a disposable spreader. The GalK cassette protocol required an extra 3 washing steps with M9 salts by pelleting and re-suspending before plating out. Selection plates were inverted once dry and incubated for 3 days at 32°C and were stored at 4°C after this time. Single colonies were chosen by colour (blue for positive selection and white for negative selection) for Rpsl and SacB methods of recombineering and colonies were grown overnight for further use. Generally 10 colonies were selected to be analysed by polymerase chain reaction (PCR), restriction digest and sequencing.

M9 salts: 6 g sodium phosphate dibasic (Na2HPO4), 3 g potassium dihydrogen phosphate (KH2PO4), 1 g ammonium chloride (NH4Cl), and 0.5 g sodium chloride (NaCl) in 1l of dH_20.

2.1.5 Polymerase Chain Reaction

Polymerase Chain Reaction (or PCR) was used to amplify specific segments of DNA such as selection cassettes for use in recombineering, or to send sections of the genome for sequencing. Reagents were used from the Expand high fidelity (HiFi) PCR system kit (Roche), unless otherwise stated. The HiFi enzyme consisted of a blend of Taq DNA polymerase and a thermostable DNA polymerase with proof-reading, and was used as

standard for the PCR of miniprep DNA, including that of selection cassettes and for the PCR of sequencing regions, with the PCR mix indicated in **Table 2.3.** The Phusion enzyme PCR kit (Thermo Fischer scientific, K0191) was used for TOPO cloning steps when recombineering US16-V5 and US17-V5 (**Section 2.1.6**), with the PCR mix seen in **Table 2.4**. Primers were designed and purchased from Sigma-Aldrich at desalted purity and were diluted to 100pM, and all primers are listed in **Appendix Table 7.1.** Primers were diluted 1/10 in dH_2O before being added to the PCR mixture. The enzymes were added last to the mix for stability.

The standard PCR program (**Table 2.5**) was run using a PCR machine (T3000 Thermocycler, Biometra). PCR programs could be adjusted according to the primers and DNA used. Annealing times could be altered dependent on the melting temperatures of the primers, and extension times could be altered depending on fragment length. Longer extension times were required for the production of longer products, with a longer PCR program required for the PCR of the larger SacB cassette as described in **Table 2.6**. A slightly altered program was also required for Phusion PCRs (**Table 2.7**).

2.1.6 TOPO cloning

TOPO cloning steps were utilised in adding the V5 tag to the C-terminus of US16 and US17 using the TOPO cloning kit (Thermo Fisher scientific) as depicted in **Section 3.1, Fig. 3.2D**. Briefly, the US16/US17 sequencing forward primer and the V5 reverse primer were used to add the V5 tag to the C-terminus of the US16/US17 gene and this was added to the TOPO vector. Once the V5 tagged US16/US17 PCR product was produced, the TOPO cloning reaction could be set up, with 0.5-4 µl fresh PCR product added to 1 µl TOPO vector, 1 µl salt solution, with water added to a total volume of 5 µl. These should be mixed gently and incubated for 5 min at room temperature (or longer for PCR products over 1 kb). A further 18 µl is added to the reaction mixture, and 2 µl of this is added to DH5α competent cells, in a cuvette and electroporated at 2.5 kV as previously described (**Section 2.1.4.3**). Immediately 250 µl of room temperature super optimal broth (SOC) medium (provided) was added and the mixture incubated in a shaking incubator at 37°C for 1 hour. 10-50 µl from each transformation was spread onto a pre-warmed selective plate and incubated overnight at 37°C. As before, 10 colonies were selected and analysed for the insert. Plasmid DNA was isolated (**Section 2.1.10**) and analyse by restriction digest (**Section 2.1.12**) and/or sequencing (**Section 2.1.13**) as normal. For clones that looked correct, the US16/US17 sequencing forward primer and the V5 forward primers could then be used to

Table 2.3: PCR mixture for HiFi polymerase PCR reaction

Ingredients	Amount
10x HiFi PCR buffer	5 µl
Dimethyl sulfoxide (DMSO) (Sigma-Aldrich, 41647)	1.5 µl
1/10 primer mix (Forward and Reverse)*	2.5 µl
Deoxynucleotide (dNTP) mix (New England Biolabs, N0447L)	1 µl
DNA	1 µl
HiFi polymerase enzyme	0.5 µl
dH_20	Make up to 50 µl

Table 2.4: PCR mixture for Phusion polymerase PCR reaction

Ingredients	Amount
10x Phusion PCR buffer	5 µL
1/10 primer mix (Forward and Reverse)	2.5 µl
Deoxynucleotide (dNTP) mix	0.5 µL
DNA	10–100 ng
Phusion enzyme	1 µL
dH_20 (sterile, distilled)	Make up to 50 µL

Table 2.5: Standard PCR program

PCR step	Temperature and time	Number of cycles
Initial denaturing	95°C for 2 min	1
Denaturing	95°C for 30 sec	35 cycles
Annealing	58°C for 30 sec	
Extension	72°C for 2 min	
Final extension	72°C for 15 min	1
Pause	Held at 4°C	1

Table 2.6: Long PCR program for amplifying the larger SacB cassette

PCR step	Temperature and time	Number of cycles
Initial denaturing	95°C for 2 min	1
Denaturing	95°C for 30 sec	34
Annealing	55°C for 30 sec	
Extension	68°C for 4 min	
Final extension	72°C for 15 min	1
Pause	Held at 4°C	1

Table 2.7: PCR program for Phusion

PCR step	Temperature and time	Number of cycles
Initial denaturing	94°C for 2 min	1
Denaturing	94°C for 1 min	25
Annealing	55°C for 1 min	
Extension	72°C for 1 min	
Final extension	72°C for 7 min	1
Pause	Held at 4°C	1

amplify the V5-tagged US16/US17 gene from the TOPO vector, whilst providing the addition of the intergenic region of homology after the V5 tag (**Figure 3.2D**). This PCR could then be used for the second step of recombineering (**Section 2.1.4**) where the US16/US17 gene disrupted by the GalK cassette could be swapped for the V5-tagged US16/US17 gene.

2.1.7 Agarose gel electrophoresis

The size of the PCR product amplified was determined by agarose gel electrophoresis. Agarose gel was made by adding 1.5% (0.75 g) HiRes standard agarose powder (AGTC Bioproducts/GeneFlow, A4-0700) to 50 ml 1x Tris-acetate-EDTA (TAE) buffer (50x, National diagnostics, EC-872) and microwaved to dissolve. Once cooled to 50-60°C (~15-20 min), 2.5 µl Ethidium bromide (Sigma) was added for DNA visualization and the solution poured into a gel mould that was taped at either end, and had a comb added to produce wells. Once solidified, the comb and tape were removed, and the gel transferred into a gel tank containing 1x TAE buffer that covers the gel. 6x DNA loading buffer was added at X1 concentration to the PCR product mixture and the samples loaded alongside 10 µl of a DNA ladder (HighRanger Plus 100kb DNA ladder, Norgen). This gel was run at a voltage of 100V for ~45 min or until the lower dye front reached the bottom of the gel to allow for separation of the product by size. The bands were visualized under UV light using a GelDoc system (Syngene) and the gel band excised if required for further processing or sequencing.

6x DNA loading buffer: 0.25% bromophenol blue (Sigma-Aldrich), 0.25% xylene cyanol FF (Sigma-Aldrich), 30% glycerol (Fisher Scientific) in dH$_2$0.

2.1.8 Purification of DNA gel fragments from agarose gel slices

UV light from a Spectroline transilluminator (model TVC-312A) was used to visualise the ethidium bromide-stained DNA bands whilst using a UV visor. DNA bands of the expected size were excised from the gel using a scalpel (Swann-Morton) and transferred into 1.5 ml eppendorfs. The gel band was then purified using the Q spin gel extraction/PCR purification kit and buffers (GeneFlow, K10040) following manufacturer's instructions. For every mg of gel that was excised, 1µl of binding buffer was added, and the gel sample heated to 50-60°C. Once the gel slice had dissolved, the solution was added to a GeneFlow column in a collection tube and left for 1 min. All spins were carried out at 13,000 rpm in a benchtop centrifuge (Biofuge fresco, Heraeus). The tubes containing the columns were centrifuged

and the flow-through discarded. 500 µl of the wash buffer was added to the column and the samples centrifuged again for 30 sec. This spin and wash step was repeated, and after the flow-through discarded, the column and tube were spun once more and the column transferred to a clean collection tube. 30µl of elution buffer was added to the column membrane for 1 min, and the DNA eluted during a final 1 min spin.

2.1.9 Measurement of DNA concentration using the Nanodrop spectrophotometer

The concentration of eluted DNA was measured using a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific). 1µl of water was used to calibrate the nanodrop machine, then 1µl of the elution buffer used as a blank. 1µl of sample DNA was then loaded and the DNA concentration was given in ng/µl.

2.1.10 Small scale purification of BAC DNA

Minipreparation or 'miniprep' allowed for the extraction and purification of BAC DNA from a small scale SW102 culture. The SW102 culture containing the appropriate BAC was grown in 5ml LB with selection antibiotics (**Section 2.1.2**) overnight at 32°C in a shaking incubator. The samples were kept and spun at room temperature unless otherwise stated. The overnight culture was pelleted at 4000 rpm for 5 min using a benchtop centrifuge (Biofuge fresco, Heraeus), the supernatant discarded and the pellet re-suspended in 250µl P1 buffer (Qiagen, 19051) and transferred to an eppendorf tube. 250µl of P2 buffer (Qiagen, 19052) was added and mixed then incubated at room temperature for 5 min. 250µl P3 buffer (Qiagen, 19053) was added and inverted to mix. This was centrifuged at 13,000 rpm for 10 min and the supernatant transferred to a new tube. The DNA was precipitated with 750 µl isopropanol and centrifuged at 13,000 rpm for 10 min at 4°C. The supernatant was removed and 500 µl 70% ethanol was added before centrifuging at 13,000 rpm for 10 min. The supernatant was removed and allowed to air-dry before being re-dissolved in 40 µl Tris EDTA.

Tris EDTA: 10mM Tris-Cl (Fisher scientific) and 1mM Ethylenediaminetetraacetic acid (EDTA) in dH$_2$0, adjusted to pH8.

2.1.11 Large scale purification of BAC DNA

Maxipreparation or 'maxiprep' allowed for the extraction and purification of BAC DNA from a large scale SW102 culture in order to get transfection-quality DNA. The SW102 culture containing the appropriate V5-tagged BAC was grown in 250 or 500 ml LB with selection antibiotic overnight at 32^oC in a sterile conical flask in a large shaking incubator (Gallenkamp). The DNA from the culture was purified using the Nucleobond BAC100 kit (Macherey-Nagel) following the instructions of their 'Low-copy plasmid purification' section under 'Maxi/BAC'. Firstly the culture was transferred to 250 ml polycarbonate centrifuge bottles (Thermo Fisher scientific, CFS-300 520C), and centrifuged at 6000 rpm for 10 min at 4^oC (Beckman Coulter centrifuge, rotor JLA 16.250). The supernatants were discarded and the pellets re-suspended in 24 ml of buffer S1. 24 ml buffer S2 was then added to lyse the cells and the mixture inverted 4-6 times and incubated for 5 min at room temperature. 24 ml buffer S3 was then added and the bottle incubated for 5 min on ice. The supernatants were passed through a funnel lined with filter paper in to a column, which was previously equilibrated with 6 ml N2 buffer. The column was allowed to empty by gravity flow and the flow through discarded. The column was then washed twice in 18 ml N3 wash buffer by gravity flow. The DNA was then eluted in to 30 ml polypropylene tubes (Thermo Fisher Scientific, 03719) using 15 ml N5 buffer which was pre-heated to 50°C for higher recovery of DNA. 11 ml room temperature isopropanol was added and the tubes centrifuged at >15000 rpm and 4°C for 30 min (Beckman Coulter centrifuge, JLA25.50 rotor). The supernatant was discarded and the pellet washed in 5 ml 70% ethanol, and centrifuged at >15000 rpm for 15 min at room temperature. The supernatant was removed and the ethanol allowed to evaporate from the pellet (~20 min). The DNA was re-suspended in 250 µl elution buffer (Tris EDTA) and left in a shaking incubator for 30-60 min to aid recovery. The eluate was then transferred to a fresh tube, and DNA concentration was measured using a Nanodrop ND1000 spectrophotometer (**Section 2.1.9**).

2.1.12 Restriction endonuclease digestion

Restriction endonuclease digests were performed by adding 8 µl of the purified miniprep DNA mixture to 1 µl of the restriction endonuclease enzyme and 1µl of corresponding buffer (**Table 2.8**). BamHI was used as standard. The samples were then incubated at 37ºC for 4 hours or overnight and separated by agarose gel electrophoresis. Restriction endonucleases digest the HCMV BAC at their specific enzyme binding sites in the genome, digesting it into a predictable pattern of bands of particular molecular weights as predicted

by CLC workbench software (CLC bio, Qiagen). **Figure 2.1** depicts the predicted patterns of the HCMV BAC (pAL1111) with BamHI, NdeI and HindIII enzymes, alongside reference molecular weight markers and the total number of bands predicted for each digest. Samples could also be compared to the equivalent digest of an unaltered HCMV BAC as a control. These patterns were used to analyse the BACs and to assess whether the recombineering protocol had caused any major off-target affects such as large deletions of the genome or major rearrangements. These changes would cause major differences in the BAC digest patterns to those predicted (**Figure 2.1**).

2.1.13 Sanger sequencing

To verify the addition of the V5 tag and to check that no other alterations had been made to the tagged gene, the tagged region of the BACs were sequenced. The region of interest (the gene and V5 tag section) was amplified by PCR from miniprep DNA using specific forward (For) and reverse (Rev) sequencing primers (**Appendix Table 7.1**). This DNA was run by gel electrophoresis and purified from the agarose gel. The DNA concentration was measured and diluted to 5 ng/μl with dH_2O. 15 μl of this mixture was added to a Eurofins sequencing tube (Mix2Seq Kits, Eurofins) with 2μl of each appropriate sequencing primer (100 pM), either the forward, reverse or internal primers of the appropriate gene (**Appendix Table 7.1**). This was then sent to Eurofins for sequencing (Eurofins Genomics sequencing department, Germany). Sequencing was analysed using CLC workbench software (CLC bio, Qiagen) through alignment to the reference sequence of the HCMV Merlin genome (NCBI NC_006273.2). Internal reverse sequencing primers proved to give the cleanest result for the area of interest but both a forward and reverse read were taken to verify that the tag had been inserted correctly. Verified clones were then maxiprepped and transfected into fibroblasts.

2.1.14 Purification of virus DNA

Viral DNA from virus stocks was purified using a QIAamp MinElute virus kit (Qiagen, 57704), with all reagents supplied unless otherwise stated, and manufacturers' instructions followed. Briefly, 25 μl protease (reconstituted in buffer AVE) was added to a 1.5 ml micro-centrifuge tube, followed by the addition of 200 μl of aliquoted virus stock within a tissue culture class II cabinet. 200 μl Buffer AL was added and the tube pulse-vortexed for 15 sec before incubation at 56°C for 15 min in a heating block. Samples were briefly centrifuged

Table 2.8: A list of the restriction endonuclease enzymes with their corresponding buffers

Enzyme	Buffer	Company
BamHI	Buffer 3.1 (or 2.1)	New England Biolabs (NEB)
HindIII	Buffer 2.1	NEB
NdeI	Buffer 2.1 (or 3.1)	NEB
BamHI	Buffer E	Promega
HindIII	Buffer E	Promega
NdeI	Buffer D	Promega

Figure 2.1: Restriction endonuclease digest patterns of HCMV BAC (pAL1111).

BamHI
51 restriction sites

HindIII
25 restriction sites

Nde1
21 restriction sites
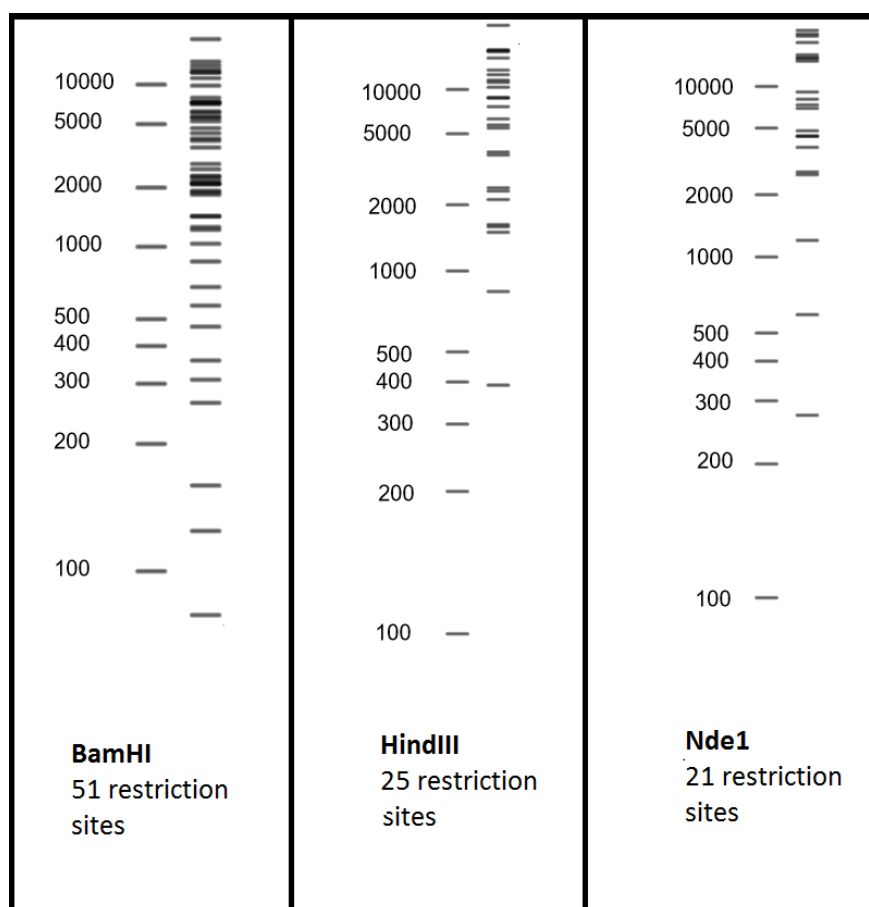
(Biofuge fresco, Heraeus) to ensure all liquid was returned to the bottom of the tube and 250 µl of ethanol (96-100%) was then added, and the tube pulse-vortexed for 15 sec. The lysate was incubated with the ethanol for 5 min at room temperature. Samples were briefly centrifuged and the lysate from step 7 carefully applied onto the QIAamp MinElute column without wetting the rim, and centrifuged at 8000 rpm for 1 min. The QIAamp MinElute column was placed in a clean 2 ml collection tube, and the filtrate discarded.

Next, 500 µl of Buffer AW1 was added, the sample centrifuged and a new collection tube used. Repeat with 500µl of Buffer AW2 and then then 500µl of ethanol (96-100%), centrifuging between each buffer addition and using a new collection tube each time, discarding the flow through. The samples were then centrifuged at 14,000 rpm for 3 min to dry the membrane completely. The column lid can be opened and the tube incubated at 56°C for 3 min to aid this process. The QIAamp MinElute column can then be placed in a clean 1.5 ml microcentrifuge tube and 50 µl of Buffer AVE applied to the centre of the membrane. The tube is incubated at room temperature for 1 min with the lid closed, and then centrifuged at 14,000 rpm for 1 min to elute.

2.1.15 Illumina full genome sequencing

Purified viral DNA was sent to our collaborators in Glasgow (Davison group, MRC-University of Glasgow Centre for Virus Research) for next generation DNA sequencing of the entire viral genome. Their analysis compared the viral genomes to the reference Merlin genome (NCBI NC_006273.2) and provided us with a list of changes, if any, other than the V5 tag addition, which are detailed in **Section 3.3, Table 3.1**.

2.2 Tissue culture

2.2.1 Established cell lines

HCMV viruses were cultured in human foreskin fibroblasts (HFFs) were provided by Dr G. Farrar (Porten Down) that were immortalised with telomerase (HF-TERTs) by Dr B. McSharry (as described in McSharry et al. (2001)). For infections with adenovirus, HF-TERTs transfected with the Coxsackie-adenovirus receptor (HF-CARS) were used (Stanton et al., 2008, Leon et al., 1998). All cell lines were negative for mycoplasma, as screened for by S. Llewelyn-Lacey using the VenorGeM Mycoplasma PCR detection kit (Biochrom AG, Germany).

2.2.2 Tissue culture media

Cells were grown in high glucose Dulbecco modified Eagle's medium (DMEM, Sigma, D5796) supplemented with 10% fetal bovine serum (FBS, Gibco, Life technologies, 10500) and 5% penicillin/streptomycin sulphate (pen/strep, Gibco, Life technologies, 15070-063), hereafter referred to as 10% DMEM. Media was pre-warmed before adding to the cells, and the cells subsequently grown at 37°C in a 5% CO_2 incubator (Thermo Fisher, BB15). DMEM with 0% or 5% FBS could also be used during experiments, 0% DMEM for infections and 5% for growing viral stocks in cell factories (**Section 2.2.4**)

2.2.3 Culturing of cell lines

All cell culture work was undertaken within a class II hood for sterility. Both HF-TERTS and HF-CARs were cultured as monolayers in 150 cm^2 Corning cell culture flasks (Sigma-Aldrich, CLS430825), referred to as T150 flasks. When T150 flasks were confluent, the fibroblasts were passaged by removing the medium, washing with ~20 ml PBS (Gibco, 14190-094) and adding 5 ml 0.05% trypsin-EDTA (x1, Thermo Fisher, 25300-054) for ~3 min until the cells detached. The trypsin was neutralized with 10% DMEM and the suspension was split 1:4 per flask for the seeding of new flasks, or were used to seed flasks or wells for experiments (as per **Table 2.9**).

2.2.4 Seeding of cells for experiments

Once the trypsinised cells had been neutralised, the number of cells could be estimated. A 100 µl sample of the suspension was loaded in to a counting chamber haemocytometer (Sigma, Z359629) and a coverslip added. Cells were counted across the 4x4 chamber grid and an average number of cells per grid calculated. This number was multiplied by 10^4 to give the average number of cells per ml of suspension. The required amount was then transferred to the flasks or wells needed for the experiment and topped up with the appropriate amount of medium (**Table 2.9)** and left overnight to settle.

Table 2.9: Flasks and plates utilised for tissue culture experiments, their uses and the number of cells required for each.

Flask/plate	Use	Cells seeded	Total amount of DMEM
T150 (150 cm^2, Corning, Sigma-Aldrich, 430825)	Growing and maintaining cells/ SILAC infections	~4x10^6 (a confluent flask will hold 18x10^6)	25 ml
T25 (25 cm^2, Corning, Sigma-Aldrich, 430639)	Transfections of BAC DNA and 72 hour infections for immunoblotting	1x10^6	7 ml (2 ml for infecting cells)
6 well plate (Thermo Fisher)	HCMV titrations	3x10^5	4 ml (1 ml for infecting cells for titrations)
12 well plate (with coverslips) (Thermo Fisher)	Fluorescence microscopy	1x10^5	2 ml (1 ml for infecting cells)
Cell factory (Thermo Fischer, 140004TS)	Growing viral stocks	Add all cells from 5 confluent T150s	500 ml whilst growing/infecting, 250 ml whilst collecting viral supernatant

2.3 Procedures involving adenovirus

2.3.1 Preparation of adenovirus stocks

Replication deficient adenovirus (RAd) stocks of US12 family V5-tagged vectors were from the Cardiff University Genebank and had been diluted in DMEM to 10^8 pfu/ml for ease of use and were stored at -70°C.

2.3.2 Infections with replication deficient adenovirus

For all infections, HF-CARs were infected the day after being seeded and were infected with virus at an MOI of 10 in 10% DMEM and incubated on a rocker at 37°C for 2 hours. The virus suspension was then removed and replaced with 10% DMEM and left to incubate for the time required for infection, with 72 hpi used as the standard time-point. Cells were then harvested (**Section 2.4.5**) for further analysis.

2.4 Procedures using HCMV

2.4.1 Transfection of fibroblasts with BAC DNA

HCMV BACS were transfected into fibroblasts using the Basic Fibroblast Nucleofection kit (Lonza, VPI-1002) to allow for replication and growth of the viruses. 1×10^6 HF-TERTs were centrifuged at 600 rpm for 10 min, the supernatant discarded and the cell pellet re-suspended in 100 µl of Nucleofector solution and transferred to a cuvette (provided). 1-5 ng/µl of each maxiprepped BAC was added to the HF-TERTs in the cuvette and transfected using program T-016 on a Nucleofector (Amaxa II/2b Device, Lonza). Transfected cells were added to pre-warmed 10% DMEM media in a T25 flask and allowed to grow overnight and replaced with fresh 10% DMEM the following day. Viral plaques became visible within 2-4 weeks, followed by a completely infected monolayer of fibroblasts within a further 2 weeks. Viral supernatant was collected from 100% infected monolayers and stored as passage 1 (P1) cultures at -70°C and used to infect subsequent cell factories.

2.4.2 Preparation of HCMV stocks

HF-TERTS were grown in a cell factory and once confluent were infected with fresh supernatant from the transfected cells from the T25 (passage 1). Once the cell factory monolayer was 100% infected (2-4 weeks later), 200-400 ml of supernatant was collected every 2 days. This was transferred into centrifuge pots (Nalgene, Sigma-Aldrich, Z353744),

and centrifuged (Beckman Coulter, rotor JLA 16.250) for 2 hours at 14,000rpm at 21°C (due to the specification of the centrifuge pots). Previously used centrifuge pots (Thermo Scientific, discontinued) were centrifuged at 35°C. Supernatants were discarded and care was taken not to let the pellets dry out. The pellets were re-suspended in 1 ml of 10% DMEM using an inoculation loop and transferred to a 15 ml falcon. The virus pot was washed with another 1 ml 10% DMEM which was also added to the falcon. These suspensions were passed 5 times through a 19 gauge needle (BD microlance 3) and 2-5 ml syringe to break up the cell pellet, and the samples centrifuged (Hereus megafuge 1.0) at 2000 rpm for 2 min. The supernatants were transferred to a new 15 ml falcon and frozen at -70°C. Once all supernatants were collected throughout the course of infection (~5/6 supernatant collections), the entire stock was thawed and pooled together in a 50 ml falcon. These were passed through a 19 gauge needle again and then centrifuged at 2000 rpm for 2 min to remove cell debris. The supernatant was then aliquoted into 0.5-1 ml aliquots in screw cap tubes with seals (VWR, 16466-054) and frozen at -70°C.

2.4.3 Titration of HCMV stocks by plaque assay

6 well plates (1 per virus) were set up with $3x10^5$ HF-TERTS per well and left to adhere overnight. 1 aliquot of the virus was thawed and serial dilutions of 10^{-6}, 10^{-7} and 10^{-8} made in 10% DMEM (10^{-5}, 10^{-6}, 10^{-7} could be used for low titre viruses). 1 ml of each dilution was added to each well in duplicate and left for 2 hours in a rocking incubator at 37°C. The virus suspension was washed off, and overlay medium (1:1 of 2% Avicel and 2x medium) was added to prevent cell-free spread. These plates were incubated at 37°C and left undisturbed for 14 days. Cells were washed 3 times with PBS (~6 ml per well), then fixed with Crystal Violet fix stain. Plaques were counted using a standard white light microscope (Nikon TMS) and titres calculated using the average number of plaques counted per dilution i.e. average of 4 plaques at a 10^{-7} dilution gives a titre of $4x10^7$ pfu/ml.

2% Avicel: 20g of Avicel powder in 1 l of dH_20.
2x medium: 250 ml dH_20, 100 ml 10x MEM, 100 ml of FBS, 30 ml of 7.5% sodium bicarbonate, 20ml of Penicillin/Streptomycin, 10 ml L-glutamine. All reagents are from Gibco, Thermo Fisher.
Crystal violet fix stain: 0.4g NaH2PO4 (sodium phosphate monobasic), 0.65g Na2HPO4 (sodium phosphate dibasic), 0.1g crystal violet, 10 ml of 37-40% formalin and 90 ml water.

2.4.4 Infections with HCMV

The day prior to infection, HF-TERTs were seeded into T25 flasks in 0% DMEM. The following day, virus was added at MOI 10 in 0% DMEM and incubated on a rocker at 37°C for 2 hours. The viral supernatant was removed and replaced with 0% DMEM and left to incubate for the time required for infection (72 hours as standard). If required, the lysosomal inhibitor Leupeptin (Leupeptin hemisulphate microbial, Calbiochem, 108975-10mg) was added to 200µM final concentration, 18 hours before the time of harvesting (at 54 hpi for a 72 hpi infection) to prevent lysosomal degradation. Cells were then harvested (**Section 2.4.5**) for further analysis.

2.4.5 Harvesting infected cells

Cells were harvested from T25 flasks at the appropriate time post infection (72hpi as standard). The supernatant was removed and the cells washed with 7ml PBS (Gibco, Thermo Fisher), then 4 ml of PBS was added and the T25 laid on ice. The cells were then scraped into the PBS using a cell scraper (Greiner Bio-one) and transferred to a chilled 15 ml falcon tube. These tubes were then centrifuged (Hereus megafuge 1.0) at 1500 rpm for 3 min at 0°C and the supernatant removed. These protein pellets could be frozen at -20°C and re-suspended with different buffers depending on protein analysis (**Section 2.5**).

2.5 Protein analysis

2.5.1 Digestion of samples with endoglycosidases

Reagents and enzymes for glycosylation analysis were from New England Biolabs unless otherwise stated. Protein samples to be analysed for glycosylation, had their harvested protein pellets re-suspended in 100 µl of 1X glycoprotein denaturing buffer (by diluting 10X glycoprotein denaturing buffer in dH_2O). The sample was then transferred to a non-stick tube (Axygen, Corning, MCT-060-L-C) and heated to 50°C for 10 min to denature the protein sample. The sample was mixed and split into 3 non-stick tubes and incubated either untreated, or with endoglycosidase H (EndoH) or peptide-N-glycosidase F (PNGase F) with the corresponding G5 or G7 buffers (now discontinued) (**Table 2.10**) at 37°C overnight. PNGase F also required the addition of NP-40 Samples were made to a total volume of 39µl.

Table 2.10: Enzymes and buffers required for studying the N-glycosylation of proteins

Sample type	Sample amount	Buffer	Enzyme
Untreated	30µl	4µl G5 buffer, 5µl dH$_2$0	-
EndoH	30µl	4µl G5 buffer, 4µl dH$_2$0	1µl EndoH
PNGase F	30µl	4µl G7 buffer, 4µl NP-40	1µl PNGase F

2.5.2 Preparation of whole cell lysates

Before further processing, all WCL protein samples required the addition of NuPage sample buffer (4X NuPage LDS sample buffer, ThermoFischer Scientific, NP0007) and 1,4-dithiothreitol (10X DTT, Acros Organics, 426380500) both added to 1X concentration. For standard protein pellets of ~100 µl, 20 µl of 4X NuPage and 13 µl of DTT were added. For samples that had undergone overnight de-glycosylation (**Section 2.5.1**), 15 µl Nupage and 6 µl DTT were added to give a 60 µl final volume.

Re-suspended pellets then underwent sonication for 25 x 1 sec pulses at 20% power (Vibra-Cell, VCX130, Sonics and Materials) to shear the DNA, and samples were transferred to non-stick 0.5 ml tubes (Axygen). Samples were heated for 10 min to 50°C in the PCR machine (T3000 Thermocycler, Biometra) for denaturation as they would smear if they were boiled at the standard 95°C due to their transmembrane domains.

2.5.3 Protein electrophoresis

Protein samples were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using 12 or 20 well NuPage 10% bis-tris precast polyacrylamide protein gels (ThermoFischer Scientific, NP0302BOX or WG1202BOX). Gels were rinsed and loaded into a gel tank and covered in running buffer. 20 µl of the prepared whole cell lysate sample was loaded as standard into each well of a gel, alongside 10 µl of a pre-stained protein marker (Invitrogen, LC5800). Loading amounts could be lowered if required for more highly expressed proteins. The samples were separated by SDS-PAGE at 180 V until the dye front reached the bottom of the gel (~70 min).

Running buffer: 50ml 20X NuPage MOPS SDS running buffer (Invitrogen, 1936381) per 1l water

2.5.4 Transfer of polyacrylamide gels to nitrocellulose membranes

Gels were removed from their casings ready for the transfer to nitrocellulose. The nitrocellulose membrane and blotting papers were pre-soaked with 2x transfer buffer and positioned in a semi-dry transblotter (Invitrogen, previously Biorad). These were ordered as such from the bottom up: thick blotting paper, nitrocellulose membrane, SDS-PAGE gel, thin blotting paper, thick blotting paper. The gel was laid flat on the nitrocellulose membrane and after stacking, air bubbles could be rolled out to ensure proper transfer.

The lid of the transblotter was secured and the transfer run for 2 hours at 10 V (ZOOM dual power supply, Invitrogen), although 1 hour at 20 V was also appropriate.

2x transfer buffer: 50ml NuPage transfer buffer (20X, Novel, life technologies, NP0006-1), 50ml methanol and 400ml dH$_2$0

2.5.5 Immunoblotting

Nitrocellulose membranes were removed from the transblotter, rinsed in distilled water, then incubated in ~15 ml antibody extender (Thermo Scientific, 32110) for 10 min, and then rinsed 7 times with ~20 ml dH$_2$0. Membranes were then washed with Tris-buffered saline with tween20 (TBST) for 1 min and then in blocking buffer (5% milk in TBST) for 1 hour. All antibodies (**Table 2.11**) were diluted into blocking buffer to minimise background staining on the blots and blots were kept on a rocker at room temperature, except for overnight when the blots were kept on a rocker at 4°C. Each primary antibody was left on the membrane overnight in ~20 ml, then washed 4-5 times with ~25 ml TBST and the secondary HRP antibody added for 1 hour in ~20 ml blocking buffer. Following 4-5 washes, SuperSignal West Pico (Thermo Fisher scientific, 10481755) was added and the antibody detected using either the ECL–Western blotting detection system (RPN 2132; Amersham) and X-ray film and developed by Xograph Imaging systems, or by using the GelDoc system (Syngene) with images taken at multiple exposure times. Using the X-ray film system, exposures were first taken at 2 min, and this was increased or decreased depending on the expression level seen. Using the GelDoc system, images were taken at 30 sec intervals up to 15 min as standard to ensure a range of exposures covered. For both systems, exposure times were decreased for high expression proteins (such as US20-V5 which often only required 5/10 secs) and increased for low expression proteins such as MICA (which often required 30-60 min). After imaging, the membrane was stripped with ~20 ml Restore stripping buffer (Thermo Fisher). The membrane was then rinsed, re-blocked and re-probed with another antibody. Anti-V5 was generally undertaken first except in the case of negative controls or MICA which could be very weak.

TBST: 29g NaCl (Fischer scientific), 20ml Tris-Cl (Fisher scientific), 5ml 10% Tween20 (Merck, 9005-64-5), 5ml 10% TritonX (Fisher BioReagents BP151-500), made up to 1l in dH$_2$0

Table 2.11: Antibodies used in immunoblotting and their concentrations

Antibody	Details	Concentration	Typical exposure time
Primary antibodies:			
V5 tag (mouse)	Stock donated by Rick Randall	1:10,000	15 min
Actin (rabbit)	Sigma Aldrich, A2066	1:2000	2 min
UL141 antibodies M550.2 and M550.3 (mouse)	In house stock (Tomasec et al., 2005)	1:10,000 each	2 min
MICA/B (mouse)	Bam01, Bamomab	1:2000	30 min
UL99/pp28 (mouse)	Clone 5C3, SC-56975, Santa cruz	1:200	2 min
UL44/ICP36 (mouse)	Clone 10d8, Virusys	1:12,800	2 min
IE1 (mouse)	Clone 8B1.2, MAB810R, Millipore	1:200	2 min
gB (mouse)	Clone 2F12, CA005, Virusys	1:2000	1 min
Secondary antibodies:			
Anti-mouse HRP (goat)	GE Healthcare	1:500	-
Anti-rabbit HRP (goat)	GE Healthcare	1:500	-

Tris CL: 4.44g/l Tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl, Fisher scientific, BP153-500) and 2.65g/l amount Tris(hydroxymethyl)aminomethane (Tris-base, Fisher scientific BPE152-1) in dH_20, adjusted to pH8

Blocking buffer: TBST with 5% milk powder (co-operative dry milk powder)

Stripping buffer: Restore stripping buffer (Thermo, 21063)

Supersignal West Pico: 1:1 mix of the luminol/enhancer and the stable peroxide buffer reagents supplied within the kit. Reagents must be mixed immediately before being added to the membrane.

2.6 Immunofluorescence

2.6.1 Preparation of coverslips

Coverslips were added to the base of 12 well plates and sterilised with 70% Ethanol for 10 min, before washing in PBS 3 times. $1x10^5$ HF-TERTS were seeded onto the sterilised coverslips in 0% DMEM, and left overnight. The next day the cells were infected with virus at MO1 10 for 2 hours in a rocking incubator and replaced with fresh DMEM before being incubated at 37^oC for 72 hours.

2.6.2 Immunostaining of coverslips

At 72 hpi cells were washed with PBS and fixed for 10 min with 2% paraformaldehyde (PFA). Cells were washed in intracellular (IC) buffer and then IC+ buffer (IC buffer with FBS and AB serum) was then added to the cells for 10-20 min to permeabilise the cells and block Fc receptors. The primary antibodies (**Table 2.12**) were diluted in IC+ buffer and added to the cells for 30-60 min at 37^oC in the rocking incubator. The cells were then washed 3 times in IC+ buffer and the secondary antibodies and DAPI nuclear stain diluted in IC+ buffer added for 30-60 min at 37^oC in the rocking incubator. Cells were washed in IC buffer twice and then fixed with 2% PFA for 10 min. Coverslips were removed from the 12-well plate using tweezers and mounted cell-side-down onto glass slides using a drop of ProLong Gold anti-fade mountant (Invitrogen). Once dried, coverslips were sealed to the glass slide using clear nail varnish, and the slide box was covered in foil and kept in the fridge when not in use.

2% paraformaldehyde (PFA): Dissolve 20g of PFA powder to 400ml PBS, adjust pH to 7-7.4 and top up to 500ml total. Filter sterilize once cooled

Table 2.12: Antibodies and dyes used in fluorescence microscopy and their

concentrations

Antibodies	Company	Concentration
Primary antibodies:		
Anti-V5 (mouse)	Stock donated by Rick Randall	1:10,000
Anti-V5 (rabbit)	Abcam (ab15828)	1:2000
Anti-Pp28 (mouse)	Clone 5C3, Santa cruz (SC-56975)	1:200
Anti-Calnexin (mouse)	Clone C8.B6, Millipore (Mab3126)	1:400
Anti-MPZL1 (rabbit)	Clone H99, Santa Cruz (SC-366775)	1:50
Anti-TGN46 (rabbit)	Abcam (ab50595)	1:200
Secondary antibodies:		
Alexa Fluor-594 (AF594) anti-mouse (goat)	Invitrogen (A11020)	1:500
AF594 anti-rabbit (goat)	Invitrogen (A-11072)	1:500
AF488 anti-mouse (goat)	Invitrogen (A-11017)	1:500
AF488 anti-rabbit (goat)	Invitrogen (A-11070)	1:500
Dyes and stains:		
DAPI nuclear stain (4',6-diamidino-2-phenylindole)	Sigma-Aldrich	1:30,000
Lysotracker dye	Red DND-99, Thermo Fisher (L7528)	1:2500

Intracellular (IC) buffer: 0.2% saponin (Sigma, S4521), 1% Bovine serum albumin (BSA, Sigma-Aldrich, A7906) and 0.05% sodium azide (Fisher scientific, CAS-26626) into PBS.

IC+ buffer: 40ml IC buffer, 5ml FBS and human AB serum (ABS)

AB serum: Serum was provided by the Welsh Blood Transfusion service and was prepared by Mihil Patel. Briefly, he span the serum for 1 hour at 4°C at 53,300 g, then filtered it through a 0.45 µm filter, and then a 0.22 µm filter. The serum was then heat inactivated for 30 min at 56°C, and stored in aliquots in the freezer at -20°C.

2.6.3 Analysis of HCMV infected cells by fluorescence microscopy

The Zeiss microscope (Axio Observer Z1) was used to capture all fluorescence microscopy images. The Zeiss microscope's ApoTome was engaged to provide optional sectioning of the cells, providing a clearer image of 1 focal plane without scattered light. A magnification of x40 with oil (Immersol 518F, Zeiss) was used as standard. The 3 colour channels used were red using AF594 (wavelength 594nm), green using AF488 (488nm wavelength) and blue using DAPI nuclear dye (461nm wavelength). The Zen2 Pro (Zeiss) software program used for analysis provided a histogram of the levels of each colour detected. The histogram was used to adjust the exposure until background levels of fluorescence were minimal, as compared to the control samples, and at a similar level amongst samples. Images were taken separately as well as merged between the 3 channels using the Zen software, and exported as tiffs.

2.7 Virion purification

140 ml of viral supernatant was collected from cell factories that were fully infected with each of the V5-tagged US12 family members. Firstly the supernatant was concentrated down to ~15 ml using Vivaflow 50 PES cassettes (Sartorius, VF05P6, 1MDa MWCO). Samples were then further concentrated using Vivaspin 20 columns (Sartorius, VS2061, 1MDa MWCO), centrifuging the columns at 2500 g for 10 min at a time until the sample was reduced to 6/7 ml.

'Heavy' and 'light' solutions (below) were made the previous day to allow the solutions to be fully dissolved. Sodium tartrate gradients were formed using the SG50 gradient maker (Hoeffer), with 5 ml of the 'heavy' solution in the back chamber and the 4 ml of the 'light' solution in the front chamber. This allowed the solutions to mix gradually in the outlet tube, forming a gradient of solution as they were carefully poured into thin-walled Ultra-

clear centrifuge tubes (Beckman Coulter, 344057), via the use of a peristaltic pump (Pump P1, Pharmacia Fine Chemicals), by continuous top-down pouring. Resulting gradients formed with 'heavy' solution at the lowest gradient areas, and the 'light' solution forming the upper areas of the gradient (**Section 5.5, Fig. 5.8**). The purified virion solution was added above this gradient, requiring 2 gradients per virus. For each ultra-centrifugation step, opposite tubes were balanced exactly by topping them up with PBS, and all spins were performed at 90465.7 x g at 20°C, using an Optima XPN-80 Ultra-centrifuge (Beckman Coulter) and the SW41 rotor. The first spin was performed for 45 min to separate the components of the viral supernatant by density over the glycerol-tartrate gradient. The virions banded in a distinct pattern (**Section 5.5, Fig. 5.8)** and were recovered using a syringe and a 19 guage needle. To remove the gradient-derived salts and other contaminants from the purified virus particles, recovered bands were added to a new centrifuge tube and washed by gradual dilution in NaPh buffer. The virions were then pelleted from the solution by ultra-centrifugation for 1 hour. The final purified virion pellet was re-suspended in NaPh buffer and NuPage LDS sample buffer to 1x (4X, ThermoFischer Scientific) and 1x DTT. Samples can now be processed by protein electrophoresis and immunoblotting as previously described (**Section 2.53-2.55**).

Na-phosphate (NaPh) buffer: Mix 19 ml of solution A and 81ml of solution B (pH 7.4).
Solution A: 0.04M (w/w) sodium-dihydrogenphosphate (sodium dihydrate) in dH_2O.
Solution B: 0.04M (w/w) disodium-hydrogenphosphate (anhydrous) in dH_2O.
 'Heavy' solution 35% Na-tartrate: 35g Na-tartate and 65g NaPh buffer
'Light' solution 15% Na-tartrate with 30% glycerol: 15g Na-tartrate, 30g glycerol and 55g NaPh buffer

3. Construction of HCMV Merlin V5-tagged US12 family members

The US12 gene family consists of a sequential tandem array of 10 genes, designated *US12* to *US21*, that are not essential for replication *in vitro* yet highly conserved in clinical isolates, thus implying the family plays an important role *in vivo* (**Section 1.8.15**). The US12 family members have been implicated in tropism and immune evasion of the virus. The most comprehensive analysis demonstrated that the US12 family act in concert to regulate the cell surface expression of a wide array of immune ligands (**Section 1.8.15.3**) (Fielding et al., 2017). Whilst this study brought an impressive insight into the functional role of the family during productive infection, the fundamental expression properties of individual US12 family members remained largely uncharacterised.

Although the laboratory in Cardiff has extensive experience in using adenovirus vectors as a rapid and efficient way to study the expression of individual HCMV genes, the previous study of adenovirus-expressed US12 family members (Dr Ceri Fielding) demonstrated that they did not tend to function well in isolation. This was validated by adenovirus (Ad) US18 and Ad US20 which didn't appear to downregulate B7-H6 or MICA on the cell surface by flow cytometry, despite having these effects in the context of infection (Fielding et al., 2014, Fielding et al., 2017). It therefore remains important to study these HCMV genes in the context of infection. Infection with HCMV also causes large cellular transcriptional and causes morphological changes in the infected cell. These morphological changes are caused by the remodeling of the host organelles, including the Golgi complex and trans-Golgi network (Das et al., 2007), resulting in the formation of the virion assembly compartment (vAC). Certain US12 family members have been demonstrated to associate with the vAC (Das and Pellett, 2007), and thus they may not interact with other cellular and viral proteins or function normally without HCMV infection. As US12 family members appear to functionally co-operate, they may therefore not work in the same way when expressed alone (Fielding et al., 2014, Fielding et al., 2017). Proteins can also undergo different post-translational modifications or folding within different expression systems. HCMV gpUL18 for example exhibited a different apparent molecular mass and different glycosylation patterns by SDS-PAGE when expressed by an adenovirus vector, a vaccinia virus vector or during a natural HCMV infection (Griffin et al., 2005). As antibodies were not available against individual US12 family members, in order to allow for the detection of US12 family proteins within the context of HCMV, US12 family members were tagged within an HCMV bacterial artificial chromosome (BAC). The HCMV Merlin strain was selected as the background virus strain as it is the most representative strain of an original clinical virus for which an infectious BAC clone was available. This BAC had been constructed by cloning in

the genome of HCMV strain Merlin at passage 5 from its original isolation and was designated BAC1111 (Stanton et al., 2010). This Merlin BAC1111 variant had the clinical wildtype sequence, except for a frameshift in *RL13* and mutation in *UL128*, which enabled efficient replication of progeny virus *in vitro*, and generated virus progeny that were genetically stable during propagation and limited passage (Stanton et al., 2010) (**Section 1.3.1**). The low copy number BAC technology was compatible with efficient genome editing, and this BAC1111 construct therefore permitted the construction and propagation of genetically stable HCMV, and provided a reproducible, characterised and sequenced source of clonal virus. BAC1111 was therefore used as the main parental BAC for studies in which genetic manipulation of the HCMV genome was required, and US12 family members were tagged within this construct. Each of the individual US12 family members was modified with a C-terminal tag to minimize any unwanted effects on protein folding or function. A V5 epitope tag was chosen due to its small size (14 amino acids), and the fact that high-affinity anti-V5 tag antibodies were available from different species commercially. It also gave superior detection when tested in parallel with the commonly used Strep tag (Dr R. Stanton, personal communication).

3.1 Epitope tagging of HCMV US12 family members within the HCMV Merlin BAC

Epitope tagging of US12 family members in the low copy Merlin BAC1111 variant was achieved using recombination-mediated genetic engineering (termed 'recombineering') (**Section 2.1.4**), allowing for seamless modification of the BAC genome. BAC1111 was transfected into *Escherichia coli* SW102 by electroporation, and these *E. coli* also contained temperature sensitive lambda (λ) red genes integrated into the genome by a defective λ prophage. These λ red recombination genes encode for 3 important enzymes exo, bet and gam, which work together to allow recombination to occur. Exo encodes a 5'-3' exonuclease that produces 3' overhangs from double-stranded DNA, mediating its annealing and recombination with homologous DNA in the BAC. Bet (or beta) is a single-stranded DNA binding protein that promotes annealing of homologous DNA by protecting the remaining 3' single strand tail and preparing the BAC DNA for homologous recombination. Together they insert the electroporated DNA into the desired target area, creating genetic recombinants. Gam encodes an inhibitor of the *E. coli* RecBCD exonuclease in order to protect the electroporated linear DNA from degradation. These genes are under the control of a temperature-sensitive λ repressor, and drive homologous recombination

but only when induced at 42°C for 15 mins (**Section 2.1.4**). This allows the prophage-containing bacteria to be incubated at 32°C at all other times to ensure that the recombinase function was efficiently suppressed to avoid unwanted recombination. Homologous recombination was therefore only enabled at set points in the recombineering protocol, first to insert the 'selection cassette' after the target gene and then to exchange this cassette for the V5 tag.

The selection system for the majority of recombineering experiments used the Rpsl (*KanR/SmS/LacZ/Rpsl*) and the SacB (*AmpR/LacZ/SacB*) cassettes (**Fig. 3.1**), both previously optimised within the laboratory and chosen because of their efficacy and ease of use (Stanton et al., 2010). The two systems require different selection media (**Section 2.1.4, Table 2.2**); the Rpsl cassette encodes for kanamycin resistance (for positive selection) and streptomycin sensitivity (for negative selection), whereas the SacB cassette encodes for ampicillin resistance (for positive selection), and sucrose sensitivity (for negative selection) (**Figure 3.2A)**. Once the selectable marker is inserted into the Merlin BAC, the bacteria acquire resistance to the respective antibiotics. The Rpsl and SacB cassettes also contain the *lac*Z gene which encodes for β-galactosidase and acts a visual selectable marker. Additon of isopropyl β-D-1-thiogalactopyranoside (IPTG) to the selection plates stimulates *lac*Z transcription, inducing the expresssion of β-galactosidase. β-galactosidase can then cleave its colourless substrate X-gal into 5-bromo-4-chloro-indoxyl, which spontaneously dimerizes and oxidizes to form a bright blue insoluble pigment (5,5'-dibromo-4,4'-dichloro-indigo), turning the bacterial colonies blue. This chromogenic assay was utilised to select for colonies that had been successful in each of the positive and negative selection steps (**Fig. 3.3**), with blue colonies containing the cassette, and white colonies lacking it. Chloramphenicol resistance was also required at each selection step, as the Merlin BAC contains a chloramphenicol resistance marker and this ensures that all *E. coli* colonies contain a copy of the BAC, avoiding false positives.

In order to achieve the insertion of the selection cassette by homologous recombination, arms of homology first had to be added to either end of the cassettes. This was achieved whilst amplifying the cassette up by polymerase chain reaction (PCR) by using primers that contained both homology to ends of the cassette, and homology to either the C-terminus of the US12 family member, or the C-terminal intergenic region (**Fig. 3.2C).** The same primers were used to PCR amplify either the Rpsl or SacB cassette as they contained a region of identical homology, and therefore only a different PCR template DNA was required. This gave the advantage of dual use of the primers, and allowed for easy
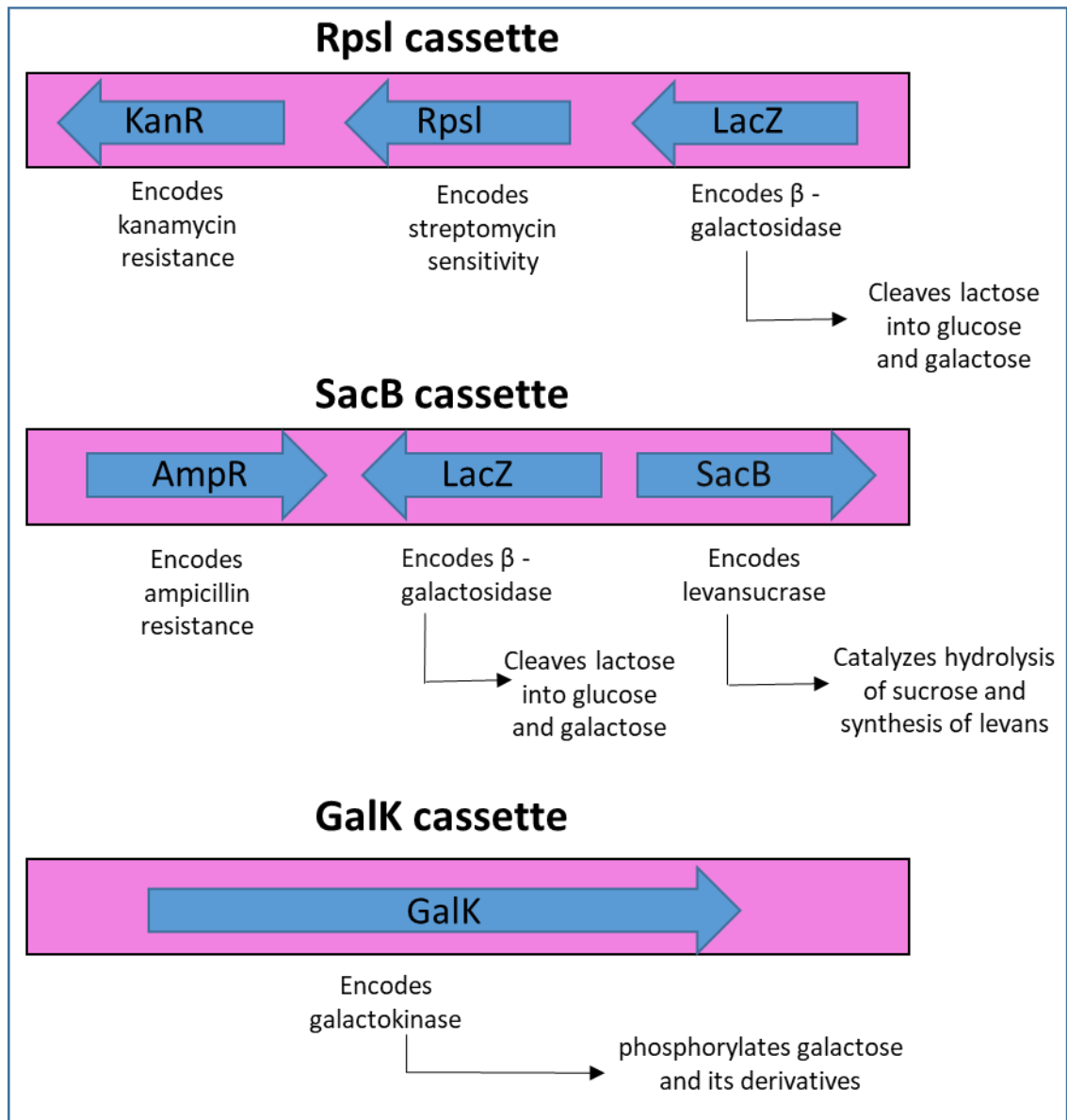
Figure 3.1: Selection cassettes used in the protocol of recombineering. Selection cassettes contain selectable markers that allow them to be selected for and against (positive and negative selection) during recombineering (**Fig. 3.2**). Selection cassettes are inserted into the region of interest (such as the C-termini of US12 family members) to allow for positive selection, and then exchanged for the intended modification (V5 tag) to allow for negative selection upon cassette removal. Rpsl and SacB are the 2 main cassettes used during the standard method of recombineering (**Fig. 3.2A**). Rpsl confers for kanamycin resistance and streptomycin sensitivity, whereas SacB confers for ampicillin resistance and sucrose sensitivity. Both also contain LacZ, a β -galactosidase which allows easy identification of colonies by blue/white screening as shown in **Fig. 3.3**. GalK is additional selection cassette that is used in the alternative method of recombineering (**Fig. 3.2B**). Positive and negative selection rounds use a different galactose carbon source, with clones containing the galK cassette able to grow in the presence of galactose, but unable to grow in the presence of its analog 2-deoxy-galactose (DOG).

Figure 3.2

C

i) Round 1

Rpsl cassette

US12 family gene

End of US12 family gene | Intergenic region

Sequence

Primer 1 (forward)

Rpsl cassette (start)

Primer 2 (reverse)

Rpsl cassette (end)

Resulting PCR product with arms of homology

Rpsl Cassette

ii) Round 2

V5 tag

US12 family gene | Rpsl cassette*

V5 tag

End of US12 family gene | Rpsl cassette* | Intergenic region

Sequence

V5 oligos (forward and reverse)

Resulting V5 tag with arms of homology insertion
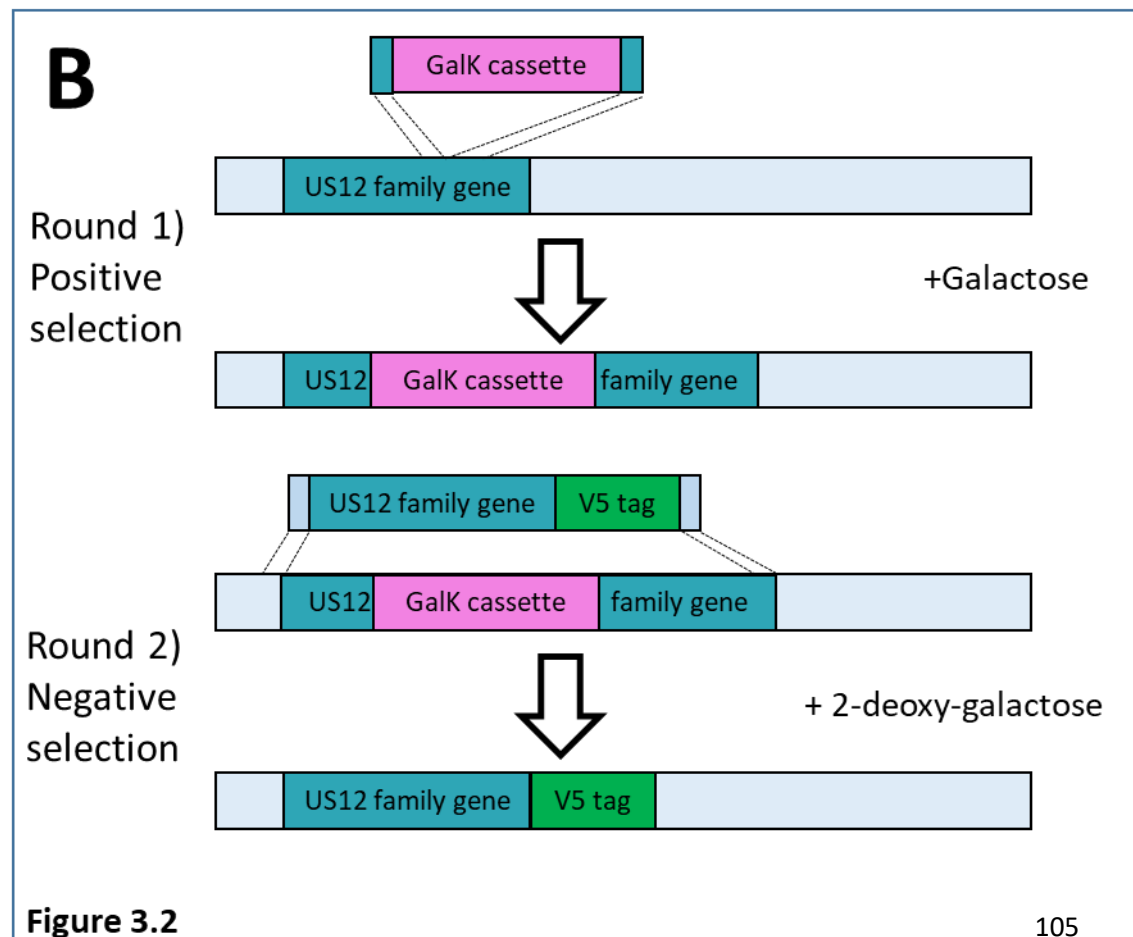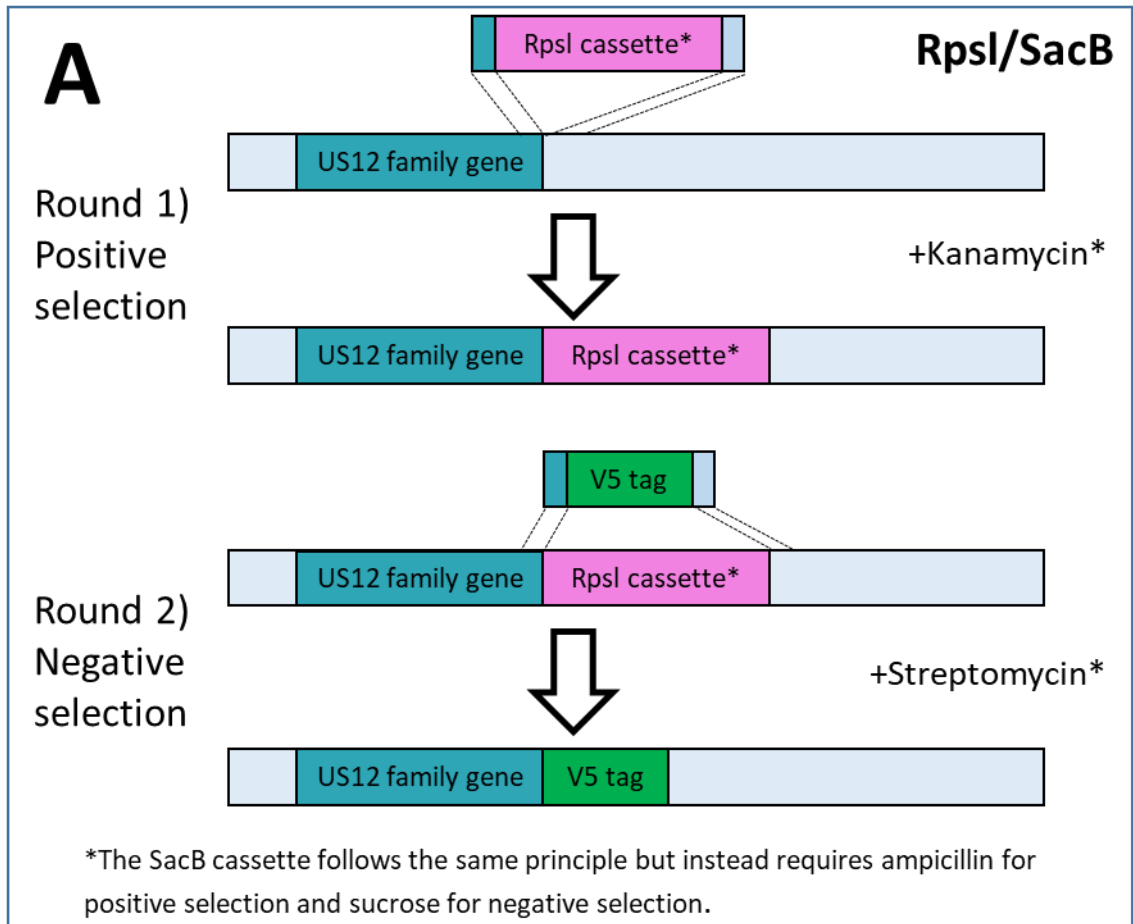
Figure 3.2 cont.

D

i) Amplify gene with V5 tag

Intergenic region | US16/US17 | Intergenic region

Primer 1 (For)

Primer 2 (V5 Rev)

Resulting US16/US17-V5 PCR
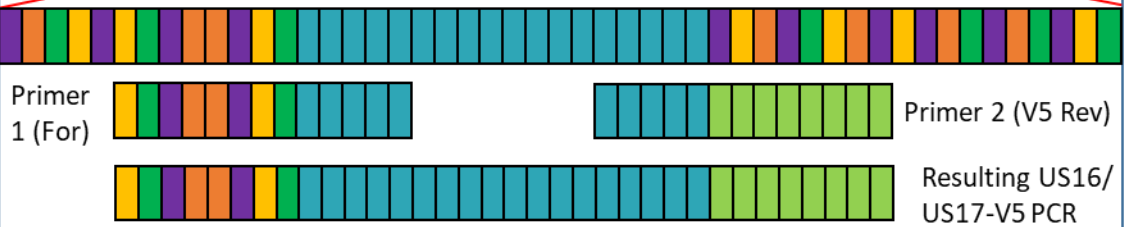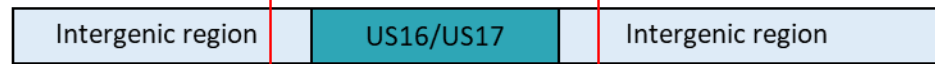
ii) Put PCR into Topo vector and amplify with arms of homology

Topo vector

Primer 1 (For)

Primer 2 (V5 For)

Resulting PCR with V5 tag and arms of homology
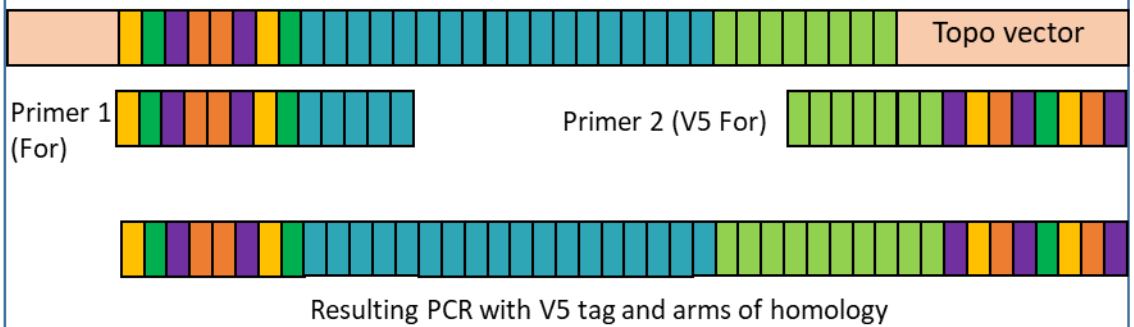
iii) Swap out the GalK-containing US16/US17 for PCRV5 tagged US16/US17 with homology arms

US16/US17 gene with GalK cassette in

US16/US17 gene | V5 tag

Resulting BAC with V5-tagged US16/US17

Figure 3.2 cont.

Key:

US12 family gene [teal block] and [striped teal block]

Selection cassettes [pink block] and [striped pink block]

V5 tag [V5 tag green block] and [striped light green block]

Topo vector [peach block]

Intergenic region [light blue block] and [multicolored block]

Homologous recombination area [dashed lines]

Figure 3.2: Recombineering strategy protocols for epitope tagging US12 family members and the strategy used to add arms of homology using PCR and primer design. Two recombineering approaches were used to V5 tag the C-terminus of US12 family members. **A)** The standard recombineering protocol utilised the Rpsl and SacB cassettes where the cassettes were inserted at the C-terminus of the US12 family member gene, and subsequently exchanged for the V5 tag. **B)** The alternative recombineering method utilised the insertion of the GalK cassette. As the clones used had selectable marker within the target gene rather than at 3' end, the cassette was exchanged for a PCR product of the complete V5-tagged gene instead. The basis of positive and negative selection by blue/white screening is explained in **Fig. 3.3**. Homologous recombination was undertaken at both stages of both protocols (rounds 1 and 2) by the addition of arms of homology to the sequences to be inserted. **C)** A more detailed representation of Rpsl and SacB recombination protocol, demonstrating how the arms of homology were added during PCR amplification. In round 1, the templates of the selection cassettes were amplified up using primers that bound to either end of the cassette and encoded for the US12 family gene C-terminus or the intergenic region, giving arms of homology at either end of the PCR product. For round 2, arms of homology were added to the V5 tag through primer/oligo design where 2 overlapping primers joined together, with arms of homology at either end. **D)** A more detailed representation of round 2 of the GalK recombination protocol, demonstrating how both the V5 tag sequence and arms of homology were added to the US12 family member gene during PCR amplification. Templates of the genes were amplified up using primers that bound to either end of the gene and encoded for the US16/US17 gene C-terminus and the V5 tag, or the intergenic region. For the addition of homology arms on the V5 terminus, the product was put into a Topo vector, and amplified up using primers that encoded for both intergenic regions of the US16/US17 gene and the V5 tag, providing arms of homology at either end of the PCR product. This entire V5-tagged gene could then undergo homologous recombination with the BACs in which the GalK was contained within the US16/US17 gene. The orange, yellow, purple and green blocks represent the sequencing pattern (A, T, G and C, in no particular order or pattern) to demonstrate the areas of homology. The pale green represents the sequence of the V5 tag, blue represents the US12 family member and pink represents the selection cassettes. Dashed lines indicate the areas of homology that lead to the homologous recombination.

LacZ
(β-galactosidase)

X-gal → X-gal product (blue)

Positive selection

Rpsl:
Kanamycin
resistance
SacB:
ampicillin
resistance

Blue colonies have
successfully
inserted the
cassette
containing LacZ

White/cream
colonies do not
contain the
cassette

Negative selection

Rpsl:
streptomycin
sensitivity
SacB:
sucrose
sensitivity

White/cream
colonies have
successfully
removed the
cassette
containing LacZ,
and hopefully
replaced it with
the V5 tag

Blue colonies still
contain the
cassette (should
rarely grow)

Figure 3.3: Blue/white screening on recombineering selection plates. This chromogenic assay of blue/white screening was used to identify which clones contained (or had removed) the Rpsl (KanR/SmS/LacZ/Rpsl) or SacB (AmpR/LacZ/SacB) selection cassettes. BAC containing colonies were grown on LB plates containing antibiotics, X-gal and isopropyl β-D-1-thiogalactopyranoside (IPTG). The antibiotics were used to select for colonies with and without the particular cassette, and the clones could be further identified by the presence or absence of *lacZ* within the cassettes which acted as a visual selectable marker. The *lacZ* gene encodes for β-galactosidase and is stimulated by (IPTG) present in the plates. This allows the X-gal in the plates to be converted into its 5-bromo-4-chloro-indoxyl product, which dimerizes to produce an insoluble blue pigment.

switching between Rpsl and SacB methods for increased efficiency of success.

Once the chosen cassette was amplified up by PCR to contain the arms of homology, the competent *E. coli* containing the HCMV BAC1111 were induced at 42°C for 15 mins. The PCR product was then electroporated into the bacteria, allowing homologous recombination to occur at the C-terminus of the US12 family gene. Bacteria were allowed to recover for 1 hour and were plated out onto selection media to select for colonies in which the homologous recombination had occurred and the selection cassette inserted. This 'positive selection' round utilised the addition of antibiotics for which the inserted cassettes contained resistance markers for (kanamycin for Rpsl and ampicillin for SacB). Blue colonies were selected for further recombineering if they grew on the positive selection plates (kanamycin/ ampicillin) but also failed to grow on the negative selection plates (streptomycin/sucrose) to avoid false positives during the second stage of recombineering.

In the second round of recombineering, the same process was followed, except the insert was the V5 tag that also contained arms of homology to the C-terminal region of the US12 family gene of interest. As the V5 tag is much smaller than the selection cassettes, instead of undergoing a PCR using primers to add the homology arms of interest to the tag, this was done solely through the addition of 2 overlapping oligonucleotides (oligos) containing both the V5 tag sequence and the homology arms (**Fig. 3.2D**). The competent *E. coli* were again induced at 42°C for 15 mins, and the oligos electroporated into the bacteria, allowing for homologous recombination to occur. The bacteria were again allowed to recover for 1 hour (or 4 hours for SacB) and were then plated onto selection media to select for colonies in which the homologous recombination had occurred and the selection cassette exchanged for the V5 tag. To avoid the growth of any colonies that still contained the selection cassettes, this 'negative selection' round utilised the addition of streptomycin for which the Rpsl cassette contained a sensitivity marker for, or sucrose which is toxic to *E. coli* still containing the SacB cassette. The loss of the selection cassette in exchange for the sequences encoding the V5 tag also included the loss of lacZ, hence white colonies were selected for (**Fig. 3.3**) and validated as below (Section **3.3**).

To achieve the tagging of members *US12*, *US13*, *US14* and *US15*, the Rpsl or the SacB cassette were firstly added into pAL1111 after the target gene, with the cassettes subsequently swapped for the V5 tag. BACs with V5-tagged *US18* and *US20* had previously been constructed and viruses generated (Fielding et al., 2014). Intermediate constructs of

US19 (*US19*-Rpsl) and US21 (*US21*-Rpsl) were produced prior to my involvement in the project (by Dr Ceri Fielding), and I completed the 2nd round of recombineering that exchanged the inserted cassettes for the V5 tag. Despite repeated attempts, the V5 tag could not be inserted at the C-termini of either *US16* or *US17* by recombineering using the standard Rpsl and SacB cassettes.

3.2 Epitope tagging of HCMV US16 and US17

It is unclear as to why recombineering with Rpsl and SacB for *US16* and *US17* was not successful. It is unlikely that the GC content of these 2 genes accounted for the incompatibility of the original cloning attempts, as the GC content of *US16* and *US17* was not dissimilar to the other members of the family (**Appendix Table 7.2)**. Perhaps the DNA sequence of these 2 genes had a secondary structure (e.g. a hairpin loop) that affected or blocked the C-terminus in some way (Nelms and Labosky, 2011). Although *US17* does have a predicted helix closer to its C-terminus than the other members of the family, in general the predicted secondary structures of *US16* and *US17* are not dissimilar to the rest of the family (data not shown). An alternative approach was adopted to epitope tag *US16* and *US17* that utilised a galactokinase (GalK) selection cassette (**Fig. 3.1**) (Warming et al., 2005). *E. coli* SW102 harbour a functional gal operon, except for the deletion of *galK*. The inclusion of *galK* on the selection cassette thereby allowed bacteria with a GalK-containing BAC to grow on minimal media with galatose as the only carbon source, allowing for positive selection (Warming et al., 2005). Galactose was replaced with its analog 2-deoxy-galactose (DOG) in the second round of recombineering. If the GalK cassette remained, DOG would be phosphorylated and lead to a toxic build-up of 2-deoxy-galactose-1-phosphate, removing any colonies that still contained the GalK cassette. Dr Eva Ruckova had previously constructed Merlin BAC clones with the Galk cassette inserted into the middle of the *US16* and *US17* target genes, albeit for a different purpose (the generation of deletion mutants). These clones were therefore used in the second round of recombineering. As the selectable marker was within the target gene rather than at the 3' end, a complete V5-tagged gene was used in recombineering rather than just the V5 tag (**Fig. 3.2B**). The PCR amplification of *US16*-V5 and *US17*-V5 used DNA primers that incorprated the V5 tag and homology arms to the *US16* and *US17* gene regions when amplifying them from the Merlin BAC1111 (**Fig. 3.2E**). These PCR products were purified and transformed into *E. coli* containing the 42°C induced Merlin BAC (*US16*-GalK or *US17*-GalK). The resulting 2nd round colonies were

grown onto the negative selection plates containing 2-deoxy-galactose (DOG) to select for those in which homologous recombination had occurred, and the GalK cassette removed. This alternative method resulted in the addition of the V5 tag to the C-terminus of both *US16* and *US17*, resulting in all 10 members of the US12 family thus being V5-tagged.

3.3 Sequencing validation of the V5 tagged genes

After the completion of both recombineering stages, the DNA purified from each V5-tagged BAC clones were PCR amplified ready for sequencing, using forward and reverse sequencing primers to amplify the region of the US12 family gene and V5 tag (**Fig. 3.4**). The size of the PCR product indicates whether is it likely that the tag had been inserted without causing unwanted deletions or additions in the gene of interest. As demonstrated by the PCR products of *US16*-V5 using forward and reverse sequencing primers (**Fig. 3.5A**), clone 13 gave a band bigger than the expected 1375 bp size, and clone 19 did not produced any strong bands of the correct size. Clones 14, 15, 16, 17, 18 and 20 however demonstrated a satisfactory quantity of PCR amplification for the region of the correct size, and were purified for sequencing.

Another step to ensure that the clones were likely not to have other unwanted changes in the viral genome, was to digest the DNA purified from each BAC clone with a restriction endonuclease such as *Bam*HI (**Fig. 3.5B**). This digest process is unlikely to detect small changes such as the addition of the tag, but instead detects major alterations in the BAC genome including unwanted recombinations that may have taken place, or large parts of the genome that may have been deleted. Therefore if the restriction endonuclease digest pattern was similar to that of the parental Merlin BAC, then the clones appeared not to have any major BAC alterations and the entire genome is likely to have remained intact. **Figure 3.5B** demonstrates acceptable restriction digest patterns from *US16*-V5 second round clones 14, 15, 16, 17, 18 and 20 that appeared suitable for sequencing. The PCR products from these clones were therefore sent for Sanger DNA sequence analysis to validate the addition of the tag. To achieve this, they were sent using forward, reverse and/or internal primers that were designed to capture the sequence of the area surrounding the addition of the tag and the entire US12 family member gene **(Fig. 3.4).** In addition to the detection of the epitope tag in the correct location, the sequence was also checked for mutations that could have been introduced by PCR or recombineering into the US12 family gene region, by comparison to the known HCMV strain Merlin US12 family
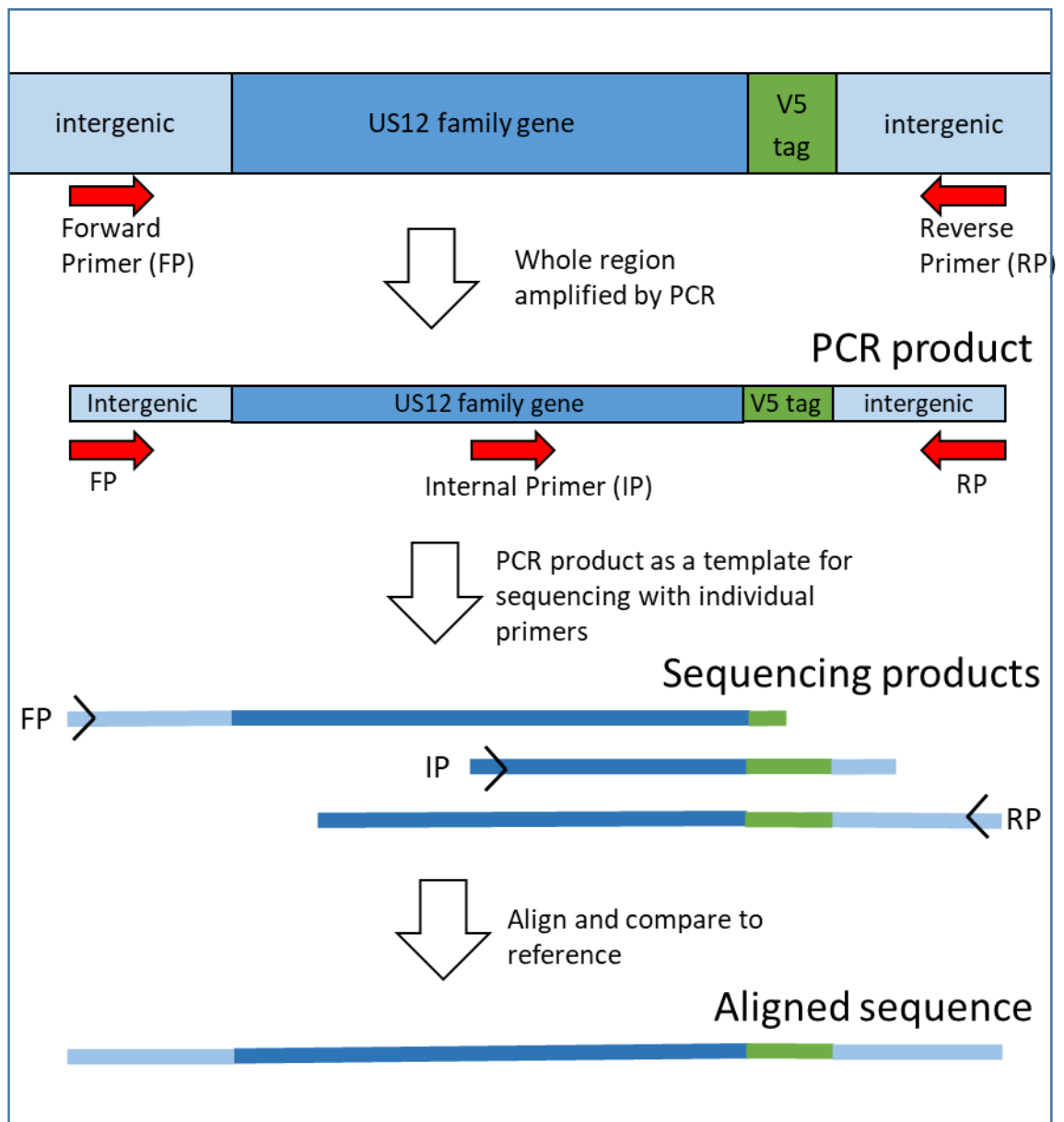
Figure 3.4: Sequencing procedure for V5-tagged US12 family members. After recombineering has taken place, V5-tagged US12 family member BACs must be verified by sequencing. Firstly, forward and reverse primers (yellow arrows) are used to amplify up the V5-tagged gene of interest from the BAC by PCR in order to send that region for sequencing. Forward, Reverse and/or Internal primers (red arrows, FP, RP, IP) were then used for the sequencing analysis of the final clones to check that the V5 tag was inserted correctly. The PCR template for each V5-tagged member was sent with each primer individually to Eurofins (Germany). The Internal primer was often used if the sequencing trace from the Forward primer wasn't clear enough to determine if the tag region had any sequencing issues. Using 2 or more primers meant that the whole region of interest could be covered by a strong sequence readout. These could be aligned and compared to the NCBI reference sequence for each US12 family member in HCMV strain Merlin.

A PCR amplification of US16-V5 region for sequencing

Clones of expected size (1375bp)
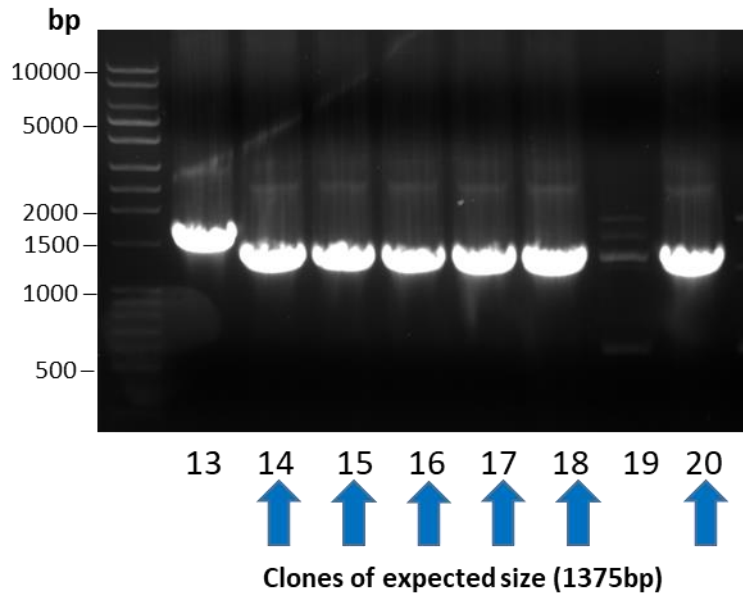
B Restriction endonuclease digest of US16-V5 BAC clones (BamHI)

HCMV 1111 digest pattern with BamHI

All clones had no major BAC alterations

#16 taken forward once sequencing confirmed
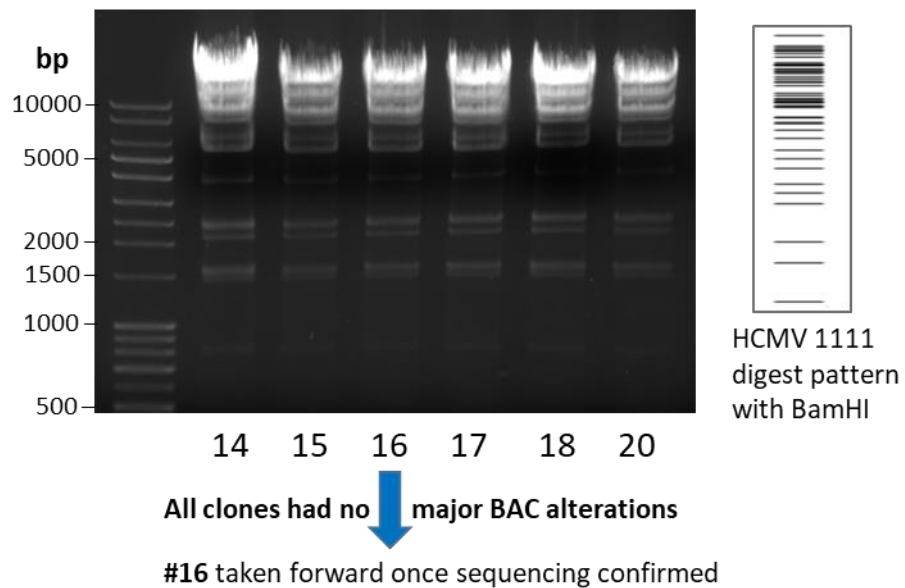
Figure 3.5: PCR and restriction endonuclease digest patterns of US16-V5 tagged clones. A) Polymerase chain reaction (PCR) of US16-V5 using both the forward and reverse US16 sequencing primers to amplify the US16 and V5 tag region for sequencing analysis. Expected band size is **1375bp**. All US16-V5 clones gave the correct size band except for clones 13 and 19. **B)** Restriction endonuclease (BamHI) digest patterns of purified US16-V5 BACs, alongside corresponding ladder pattern from BAC pAL1111 parental strain (formulated by CLC Main Workbench 7 software). All clones had a similar digest banding pattern to the parental strain suggesting that no major BAC alterations had occurred during recombineering. Multiple clones were sequenced and #16 was subsequently verified and taken forward into the next stages.

Figure 3.6

US16 forward primer   **US15 gene**

Reference
GGGTCTCTTGGATCCCGGGGGGCACGTAGATGACCGTGCCATCGGTGGGTACTTGAAACCCTTTTTCTCTTCTCATGGTGCG

Consensus
CTTTTTCTCTTCTCATGGTGCG

For primer
▯ACT▯▯GT▯T▯▯CTTTTTCTCTTCTCATGGTGCG

Rev primer

V5 forward primer

Reference
CTGCGTTCTCTGGAAACGGCTGCTCTGTCCGAAAACCAGTTCCGAACGAAAATttacgtagaatcaagacctaggagcgggtt

Consensus
CTGCGTTCTCTGGAAACGGCTGCTCTGTCCGAAAACCAGTTCCGAACGAAAATTTACGTAGAATCAAGACCTAGGAGCGGGTT

For primer
CTGCGTTCTCTGGAAACGGCTGCTCTGTCCGAAAACCAGTTCCGAACGAAAATTTACGTAGAATCAAGACCTAGGAGCGGGTT

Rev primer

V5 reverse primer

Reference
agggattggcttaccagcgctGGGCGAGAGGGTGGACAACGGCGTTGACGACGAAGCATGGGACAGGTCGTTCGGCGTTAACG

Consensus
AGGGATTGGCTTACCAGCGCTGGGCGAGAGGGTGGACAACGGCGTTGACGACGAAGCATGGGACAGGTCGTTCGGCGTTAACG

For primer
AGGGATTGGCTTACCAGCGCTGGGCGAGAGGGTGGACAACGGCGTTGACGACGAAGCATGGGACAGGTCGTTCGGCGTTAACG

Rev primer

US16 gene

Reference
TCATCGCGTCGGACGACGGTAGTTCTAAGAGACGTAGATCGCTCAGCAGGTCCTGACAGTTGCGGATTCGCAAGATCAGAAAA

Consensus
TCATCGCGTCGGACGACGGTAGTTCTAAGAGACGTAGATCGCTCAGCAGGTCCTGACAGTTGCGGATTCGCAAGATCAGAAAA

For primer
TCATCGCGTCGGACGACGGTAGTTCTAAGAGACGTAGATCGCTCAGCAGGTCCTGACAGTTGCGGATTCGCAAGATCAGAAAA

Rev primer
GATCGCTCAGCAGGTCCTGACAGTTGCGGATTCGCAAGATCAGAAAA

Figure 3.6 cont.

Figure 3.6 cont.

US16 gene

Reference
CTCCTGGCAGGCACACGTGACACCTAGCGCCGGGATGTGCGCCAGGAGGCCGGCGAGTAATAGCACCAGCTGTCGGAACGGAC

Consensus
CTCCTGGCAGGCACACGTGACACCTAGCGCCGGGATGTGCGCCAGGAGGCCGGCGAGTAATAGCACCAGCTGTCGGAACGGAC

For primer
CTCCTGGCAGGCACACGTGACACCTAGCGCCGGGATGTGCGCCAGGAGGCCGGCGAGTAATAGCACCAGCTGTCGGAACGGAC

Rev primer
CTCCTGGCAGGCACACGTGACACCTAGCGCCGGGATGTGCGCCAGGAGGCCGGCGAGTAATAGCACCAGCTGTCGGAACGGAC

US16 gene

Reference
GACGGCAGCGCGGGTGCCGGTTTCGCTGAGCGAGAACCGGTCGCTCATAGCGGAAATACACGAAGAGCGCGGAGGCCACAGGC

Consensus
GACGGCAGCGCGGGTGCCGGTTTCGCTGAGCGAGAACCGGTCGCTCATAGCGGAAATACACGAAGAGCGCGGAGGCCACAGGC

For primer
GACGGCAGCGCGGGTGCCGGTTTCGCTGAGCGAGAACCGGTCGCTCATAGCGGAAATACACGAAGAGCGCGGAGGCCACAGGC

Rev primer
GACGGCAGCGCGGGTGCCGGTTTCGCTGAGCGAGAACCGGTCGCTCATAGCGGAAATACACGAAGAGCGCGGAGGCCACAGGC

US16 gene

Reference
ACCAGGAGGAGCACCTCGGGCGCCCAGACAACGTGACAAGGAAAGCCCGGACGCGACTTGAGAGTCGCTGTAGGGAAGACCAG

Consensus
ACCAGGAGGAGCACCTCGGGCGCCCAGACAACGTGACAAGGAAAGCCCGGACGCGACTTGAGAGTCGCTGTAGGGAAGACCAG

For primer
ACCAGGAGGAGCACCTCGGGCGCCCAGACAACGTGACAAGGAAAGCCCGGACGCGACTTGAGAGTCGCTGTAGGGAAGACCAG

Rev primer
ACCAGGAGGAGCACCTCGGGCGCCCAGACAACGTGACAAGGAAAGCCCGGACGCGACTTGAGAGTCGCTGTAGGGAAGACCAG

US16 gene

Reference
AGAGAAGCTACCCAAGACGGCCACCGCCGTGGAGATTTGGAAGAGGAGCAAGCCGGCGATTCGGACGACAACCTCGAAGCGAT

Consensus
AGAGAAGCTACCCAAGACGGCCACCGCCGTGGAGATTTGGAAGAGGAGCAAGCCGGCGATTCGGACGACAACCTCGAAGCGAT

For primer
AGAGAAGCTACCCAAGACGGCCACCGCCGTGGAGATTTGGAAGAGGAGCAAGCCGGCGATTCGGACGACAACCTCGAAGCGAT

Rev primer
AGAGAAGCTACCCAAGACGGCCACCGCCGTGGAGATTTGGAAGAGGAGCAAGCCGGCGATTCGGACGACAACCTCGAAGCGAT

117

Figure 3.6 cont.

Key:

CGTAAGCT = Sequence from Forward (For) primer (green)

CGTAAGCT = Sequence from Reverse (Rev) primer (red)

➡ = Primers   ➡ = Gene

AAGGCCCACCGTGCTGTAGAA
= Consensus sequence

TCTAACAGGCTTGATTGGTGG
= Nucleotide trace (accurate)

TGAGAACGTTAGCCAGAAAGA
= Nucleotide trace (inaccurate)

Figure 3.6: Sanger sequencing alignment example of a HCMV BAC containing US16-V5. V5-tagged BACs were sent for Sanger sequencing analysis (Eurofins, Germany) individually with both forward and reverse sequencing primers and/or the internal reverse primer of the V5-tagged US12 family member, as detailed in **Appendix Table 7.1.** Both sequences and nucleotide trace patterns for provided for each 'read' of the sequence sent. The 2 sequencing reads (with trace) from the forward and reverse sequencing primers were aligned with the reference Merlin sequence (with added V5 tag sequence) in the CLC Main Workbench version 7.6 **(**Qiagen, Denmark, www.clcbio.com)**.** The reference sequence used was the HCMV Merlin NCBI ReqSeq NC_006273.2, with the sequence of the V5 tag added at the C-terminus of each US12 family member in place of the stop codon. The example here shows the sequencing reads for US16-V5, so US16 forward and reverse primers were used. The consensus sequence alignment (shown in pink) is the overall agreeing consensus of the 2 sequencing reads in comparison with the reference sequence. Conflicts are highlighted, but are only relevant if the trace is well defined at that area (accurate), and are disregarded if the trace is missing, low or overlapping (inaccurate) or if the peaks of the trace from the opposite read are clearer and therefore more accurate.

member DNA sequences (NCBI). An example of this comparison output can be seen for *US16*-V5 in **Figure 3.6**. The regional DNA sequencing analysis confirmed that each US12 family member gene had the addition of the V5 tag sequence at the C-terminus of the gene, and that no other amino acid changed had occured within the genes (**Table 3.1**).

The *E. coli* bacterial cultures containing the validated V5-tagged US12 family member BACs were then cultured on a large scale and underwent BAC maxiprep (**Section 2.1.11**) to purify the Merlin BAC DNA. This purified BAC DNA was then transfected into fibroblasts (HF-TERTs) using Lonza's nucleofector technology, leading to the generation of plaques of HCMV-infected cells within 2-3 weeks that were then expanded to produce a virus stock within 4-5 weeks (**Section 2.4.1**). HCMV viral DNA was extracted from stocks of each V5-tagged virus, and sent to collaborators (Davison group, MRC University of Glasgow Centre for Virus Research) for whole genome sequencing to confirm that no unwanted changes had taken place in the rest of the genome. All viral genome sequences came back with the anticipated C-terminal tags and were otherwise identical to the parental HCMV Merlin genome, or at worst contained single nucleotide deletions within non-coding regions of the genome (**Table 3.1**). Such mutations should not affect the overall function of the virus or the genes of interest. Once validated, the V5-tagged HCMV Merlin BAC was given a virus identifier number (e.g. RCMV2314) for easy documentation within the laboratory and the clone number recorded (e.g. clone 3C) (**Appendix table 7.3**). Multiple stocks of each V5-tagged virus were grown as required, either by transfecting viral BACs as above, or by infection with a passage 1 virus to avoid the risk of mutation.

3.4 Validation of V5-tagged US12 family protein expression by immunoblotting

To assess whether each V5-tagged US12 family member gene produced its respective V5-tagged protein *in vitro*, whole cell lysates were collected from infected fibroblasts at 72 hpi and were subjected to SDS-PAGE and analysed by immunoblotting with a V5 specific antibody. Each of the V5-tagged US12 family member proteins (US12 to US21) were detected by the V5 antibody by immunoblot (**Fig. 3.7**), further validating the V5 tagging of these proteins. Prior to SDS-PAGE, it was necessary to denature the proteins at a lower temperature than normal (50°C instead of the standard 95°C), to avoid the formation of high molecular complexes and other issues related to their 7TM nature.

Table 3.1: The US12 family member sequencing results from the Sanger sequencing of the tagged gene region by Eurofins and the Illumina genome sequencing of the viral DNA stock by our collaborators (MRC- University of Glasgow Centre for Virus Research)

Gene	Region sequencing (Eurofins, Germany)	Viral genome sequencing (Illumina)
US12-V5	V5 tag addition confirmed, and no mutations within the sequenced region	1 nt deletion in US34A-TRS1 intergenic region (G tract)
US13-V5	V5 tag addition confirmed, and no mutations within the sequenced region	No unexpected changes
US14-V5	V5 tag addition confirmed, and no mutations within the sequenced region	1 nt deletion in RNA4.9, 1 nt insertion in US13-US14 intergenic region
US15-V5	V5 tag addition confirmed, and no mutations within the sequenced region	No unexpected changes
US16-V5	Gene and V5 tag addition confirmed, and no mutations within the sequenced region	No unexpected changes
US17-V5	Gene and V5 tag addition confirmed, and no mutations within the sequenced region	No unexpected changes
US18-V5	Tagged previously (Dr Ceri Fielding), V5 tag addition confirmed, and no mutations within the sequenced region	Sequenced previously (Dr Ceri Fielding), no unexpected changes
US19-V5	V5 tag addition confirmed, and no mutations within the sequenced region	No unexpected changes
US20-V5	Tagged previously (Dr Ceri Fielding), V5 tag addition confirmed, and no mutations within the sequenced region	Sequenced previously (Dr Ceri Fielding), no unexpected changes
US21-V5	V5 tag addition confirmed, and no mutations within the sequenced region	No unexpected changes
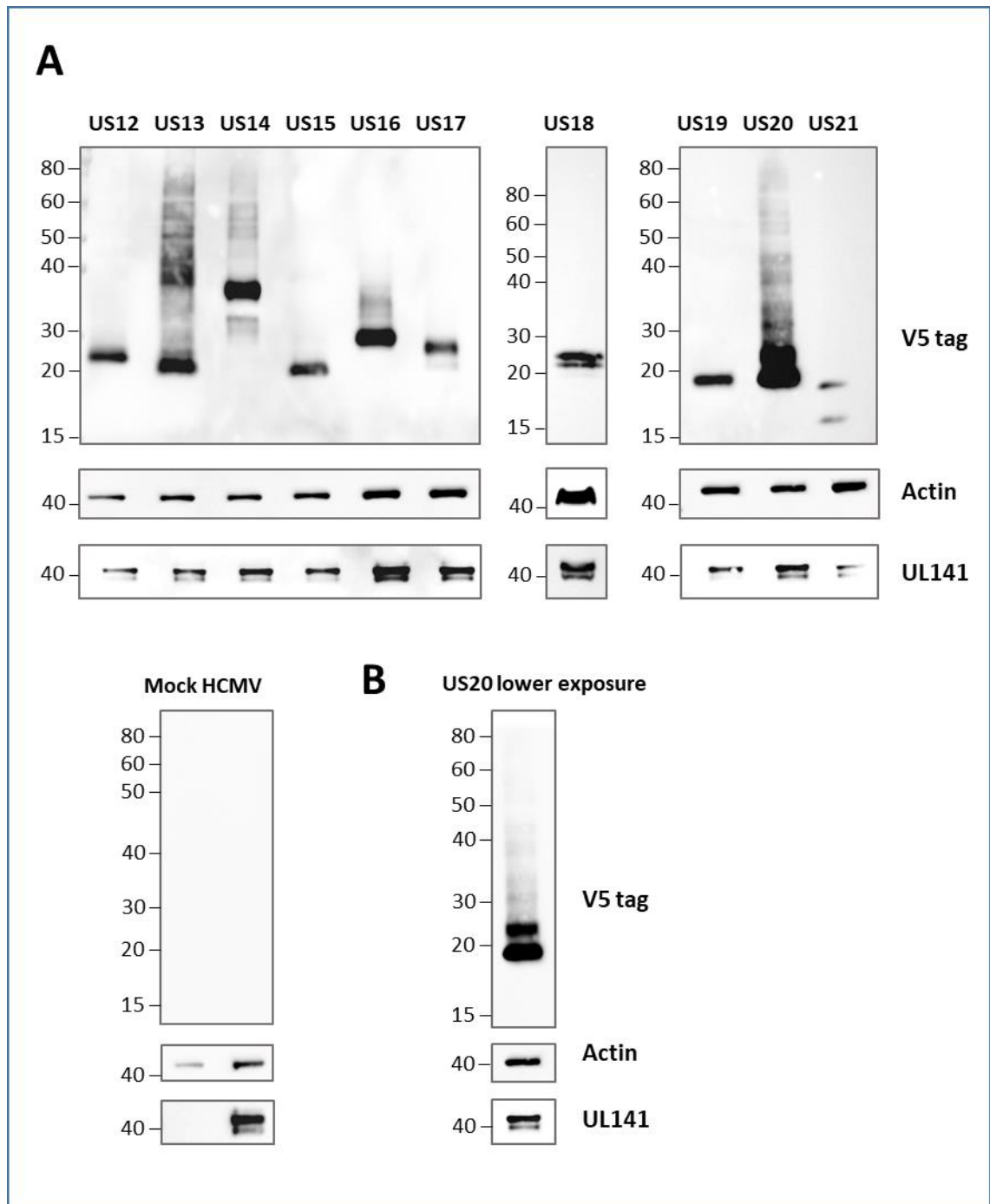
Figure 3.7: Validation of V5-tagged US12 family protein expression by immunoblotting
Fibroblasts were infected with HCMV Merlin encoding V5 tagged US12 genes at an MOI of 10. Samples were left untreated and were harvested at 72h.p.i. Whole cell lysates were extracted and proteins detected using an anti-V5 antibody (mouse) and a rabbit anti-mouse HRP secondary antibody. Samples were re-probed with an anti-actin antibody as a loading control & UL141 as a positive control for viral infection. Exposures captured using the GelDoc (Syngene) system. Mock and HCMV 1111 were run on a separate gel below. **A)** All V5-tagged US12 family members. All samples were from the same experiment, except for US18-V5 which had inefficient infection of fibroblasts in the initial experiment so was repeated separately. **B)** Lower exposure image of US20-V5 from A for clearer observation of doublet band. All samples are representative of at least 3 independent repeats.

Members pUS12-V5, pUS14-V5, pUS15-V5, pUS16-V5 and pUS19-V5 were detected as singlets (a single dominant protein species), and pUS17-V5, pUS18-V5, pUS20-V5 and pUS21-V5 (**Fig 3.7**) were detected as doublets (two dominant protein species), whereas pUS13 exhibited a smeared ladder effect. With pUS20-V5 in this and subsequent immunoblots, a shorter exposure time was required in order to reveal the doublet band (**Fig 3.7B**).

The main protein forms of each family member all fell within the 17-38 kDa range. pUS12-V5 was detected here at ~24 kDa, and pUS13-V5 exhibited a band at 21 kDa with a smeared ladder effect above reaching to ~80 kDa. pUS14-V5 was detected at ~37 kDa, pUS15-V5 at 20 kDa, and pUS16-V5 at 28/29 kDa. pUS17-V5 was detected with a main band at ~26 kDa and a fainter band at 22 kDa, with pUS18-V5 detected as a doublet of ~22 and 25 kDa, and pUS19-V5 detected at ~19 kDa. pUS20-V5 was detected as a doublet of ~19 and 24 kDa, and pUS21-V5 was detected as a doublet of ~16 and 19 kDa (**Fig. 3.7**). The molecular weights of US12 family members across this and subsequent experiments are detailed in **Chapter 4, Table 4.5**.

Despite US12 family members typically having 1 or 2 dominant protein species, other lower abundance protein species could also be detected. **Figure 3.7** demonstrated this to a small extent, however the presence of these additional protein forms were more clearly visualised at longer exposure times as demonstrated in **Figure 3.8**. pUS12-V5, pUS18-V5 and pUS19-V5 were the only members that did not appear to have any additional protein forms. Certain members, such as pUS13-V5, still showed complex migration patterns despite denaturing at 50°C. This most likely relates to the hydrophobic nature of the US12 family, but could also be caused by post translational processing (such as glycosylation, ubiquitination) or proteolysis. The relative migration of the US12 family proteins and their post-translational modifications will be investigated and addressed in more detail (**Section 4.4**) and the implications discussed further in **Section 6.1.**

At 72 hpi, the expression levels of US12 family members were clearly variable, as estimated by the strength of signal from the V5 tag. pUS20-V5 appeared to exhibit relatively high levels of expression and pUS21-V5 had relatively low expression levels compared to other family members (**Fig. 3.7**). For example, when actin levels were taken into account, the detected protein level of pUS20-V5 was estimated to be 33 times higher than that of pUS21-V5, 5-8 times higher than pUS15-V5 and pUS19-V5, and 3-4 times higher than pUS12-V5 or pUS16-V5 by densitometry (**Table 3.2**). These differences in expression levels

A
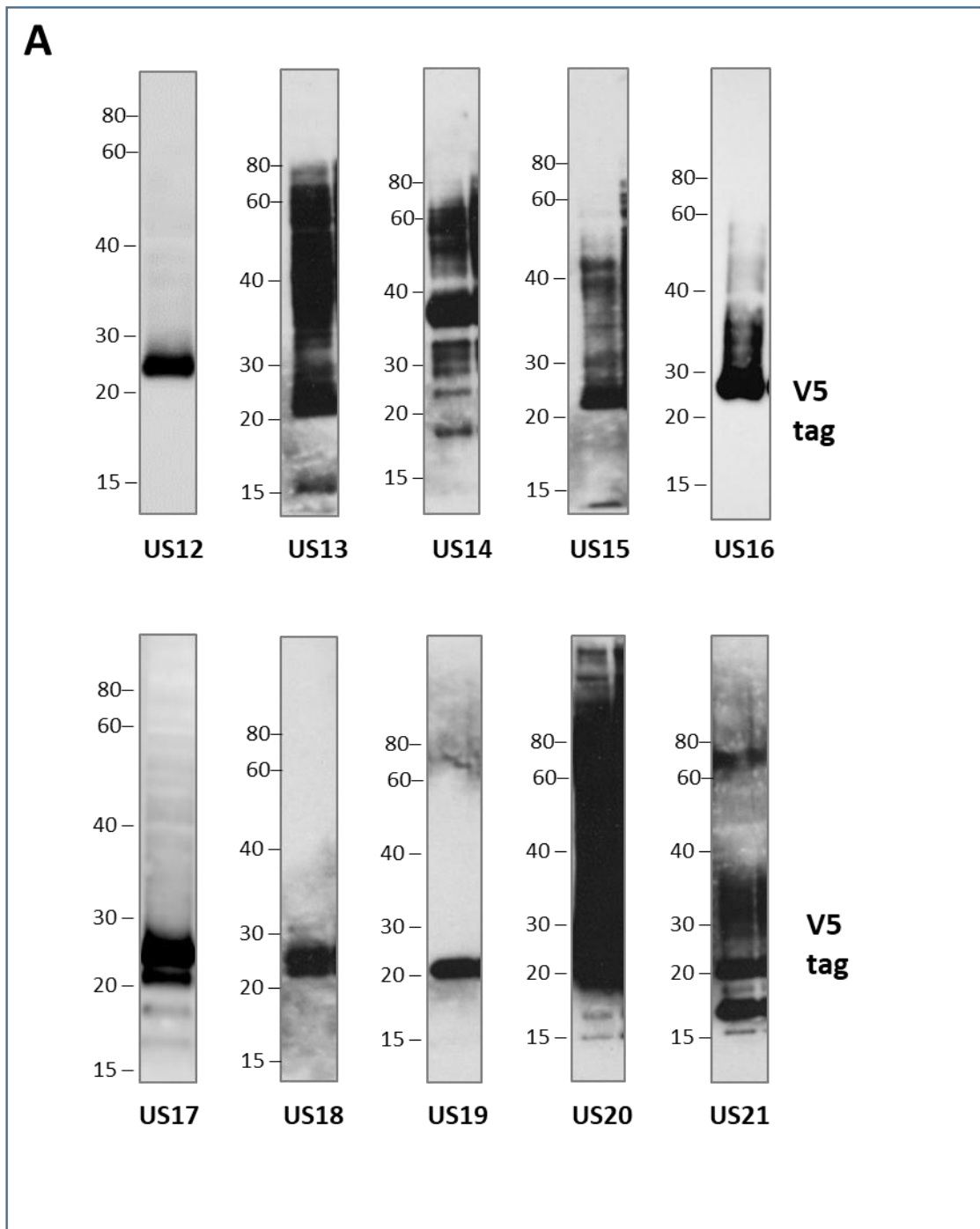
Figure 3.8: Detection of V5-tagged US12 family member proteins, at longer exposure times to visualise the extra bands and smears/ladders. At longer exposure lengths, extra bands can be detected through immunoblotting of multiple US12 family members. These are additional protein forms to the main dominant protein species typically detected, and are not consistently seen as they appear to have much lower relative expression to the main protein species. These high exposure images have been taken from across multiple experiments to find the highest available exposure for each individual protein. All samples are high exposures of untreated WCL samples at 72 hpi that have been probed by immunoblot with an anti-V5 antibody. Further protein forms were additionally detected under different treatment of the samples, including under leupeptin treatment (see **Chapter 5, Fig. 5.1**).

Table 3.2: Densitometry calculations to estimate the differences in protein expression levels between US12 family members

Protein		Area1*	Area2*	Comparative area % ◊	Comments
Figure 3.7	US12-V5	5950.640	16234.2	29.47232	3.39 times lower than pUS20
	US13-V5	35273.337	60749.24	110.287	10% higher than pUS20
	US14-V5	20277.534	36583.66	66.41565	2/3 of pUS20
	US15-V5	6631.660	10510.77	19.08173	5.24 times lower than pUS20
	US16-V5	14830.409	15196.2	27.58788	3.625 times lower than pUS20
	US17-V5	4009.569	4043.033	7.339908	13.64 x lower than pUS20
	US18-V5	†	†	†	†
	US19-V5	6157.669	6802.255	12.34913	8.097 times lower than pUS20
	US20-V5	41741.546	55082.88	100	
	US21-V5	1681.568	1681.568	3.052796	33 times less than pUS20
Figure 3.8	US18-V5	4415.225	4623.272	4.93387	20.3 times less than pUS20
	US19-V5	13673.832	16067.96	17.14742808	5.83 times less than pUS20 & 3.48 times more than pUS18
	US20-V5	93704.768	93704.77	100	

* Densitometry plots were made for each lane of protein, and the area under the peak calculated using ImageJ (NIH, https://imagej.nih.gov/ij/). Area1 is the raw area for each protein, and area2 is the area for each once the different actin protein expresssion levels were taken into account

† US18 was not present in this immunoblot so couldn't be directly compared

◊ Area % in comparison with pUS20-V5
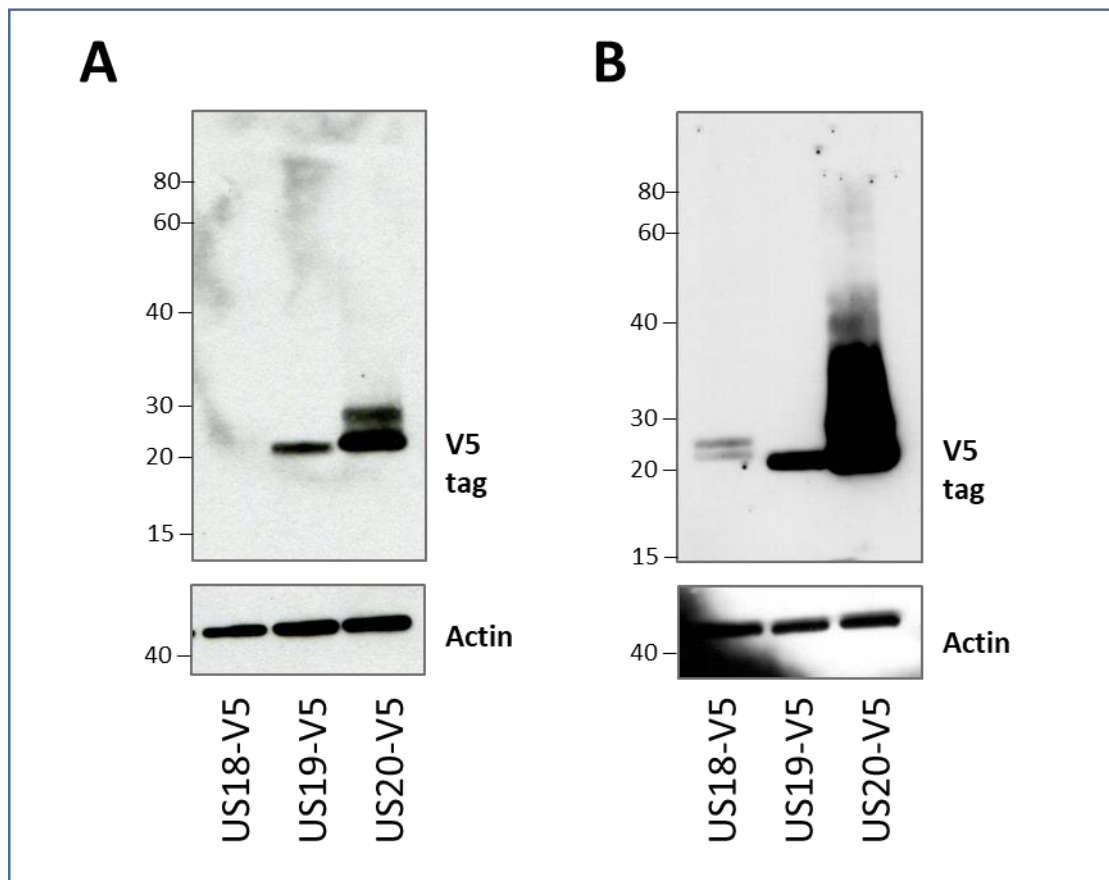
Figure 3.9: The consistent varying levels of protein expression between US12 family members as demonstrated by US18, US19 and US20. Fibroblasts were mock infected or infected with HCMV Merlin, or HCMVs encoding V5 tagged US12 genes at an MOI of 10. Samples were left untreated and harvested at 72 h.p.i. Proteins were extracted and detected using an anti-V5 antibody and an HRP secondary antibody. All samples were run with anti-actin used as a loading control. These immunoblots are just 2 examples (A and B) of the consistent difference in expression between US18, US19 and US20. Samples are representative of at least 4 independent repeats.

across the family remained consistent and could also be seen in subsequent blots (**Chapter 4**); further exemplified for pUS18, pUS19 and pUS20 in **Figure 3.9**. In this experiment, densitometry calculations indicated that pUS19-V5 was detected at ~3.5 times the amount of pUS18-V5, and that pUS20-V5 was detected at 20 times the amount of pUS18-V5, when taking actin levels into account (**Table 3.2**).

The detection of different US12 family members therefore often required different exposure times. Although the use of the same epitope tag for each US12 family member permits a comparison, the V5 tag may not be equally accessible to the antibody depending on the tertriary structure of the re-folded protein. It should also be borne in mind that post-translational processing, such as C-terminal trimming, could result in the tag being released while the protein remains functional. Nevertheless, it is reassuring to be able to detect, track and characterise expression of US12 family members during experimentation. The detected levels of each tagged protein will be influenced by the efficiency of transcription, translation and degradation and, in most cases, is likely to be an accurate reflection of expression by HCMV.

All 10 members of the US12 family were readily detected by western blot using the V5 tag, suggesting that they are likely to be functional in human fibroblasts, and demonstrates that they are all expressed in the context of productive HCMV infection. This is important as pUS13, pUS16, pUS19 and pUS21 were not detected in a comprehensive proteomic analyses of HCMV gene expression (Weekes et al., 2014). This indicates that immunoblotting using V5 tag constructs can be more sensitive than mass spectrometry. In addition to demonstrating enhanced sensitivity, the result is also helpful in interpreting the systematic proteomic analysis of individual US12 family mutants in which *US13*, *US16* and *US19* were all shown to be required to suppress expression of cell surface immune ligands (Fielding et al., 2017). Having assigned functions to these genes, it is reassuring to now demonstrate that they are expressed and can be detected.

4 Characterising expression of HCMV US12 family members

Having inserted a V5 tag at the C-terminus of each individual US12 gene family member, my intention now was to use this resource to characterise US12 family expression in human fibroblasts. Limited characterisation of individual US12 family members had been previously elucidated from focussed studies on US14, US17, US18, US19 and US20 (Guo, 1993, Das and Pellett, 2007, Das et al., 2006). A more comprehensive view was also ascertained from a high definition proteomics study that studied the expression of all quantifiable HCMV genes, including some US12 family members, and clustered them based on the similarity of their temporal profiles (Tp1 to Tp5) (Weekes et al., 2014)(**Section 1.4.1.4**). In addition to informing directly on the expression of US12 gene family members, it was intended that the study would also provide an independent assessment on the value of the proteomics analysis for this gene set. Utilising immunoblotting to study this gene family aimed to validate and extend the findings on US12 family expression by Weekes et al. (2014), and to characterise *de novo* expression (including kinetics) of the HCMV strain Merlin US12 family members not detected by proteomics.

Proteins can also be subject to a variety of post-translational modifications including ubiquitination, SUMOylation, glycosylation, phosphorylation and methylation that can impact on their properties. For example, glycosylation is known to affect folding, stability, activity, distribution, function or impair immune detection by cloaking antigenic domains (Wagner et al., 2008, Rudd et al., 2001, Helenius, 1994). Glycans can be attached either to a nitrogen residue in asparagine (N-linked) or a hydroxyl group of serine, threonine or hydroxylysine (O-linked)(Rudd et al., 2001); with N-linked (N-) glycosylation being the most prevalent. N-glycosylation sites often interact with chaperones to stabilise intermediates and enable proper folding and assembly in the endoplasmic reticulum (ER) prior to the protein's transfer to the Golgi apparatus (Rudd et al., 2001, Trombetta and Helenius, 1998, Helenius, 1994). During a further maturation process through the Golgi apparatus, involving the removal of the α1-3 and α1-6 mannose units, the N-glycosylated protein becomes resistant to endoglycosidase H (EndoH) cleavage (Freeze and Kranz, 2010). EndoH can therefore normally only remove N-glycosylation from proteins retained within the ER, whereas Peptide-N-Glycosidase F (PNGase F) removes virtually all N-glycans from a protein. Many HCMV proteins are known or predicted to be N-glycosylated, including gpUL18, gpUL141, gB and the majority of the RL11 family (Gabaev et al., 2011, Stanton et al., 2010, Griffin et al., 2005, Cochrane, 2009). Although there is no evidence that HCMV encodes its own glycosyltransferases, the virus does up-regulate the transcription of host glycosyltransferases (Cebulla et al., 2000).

Our current understanding of proteins expressed by US12 family members is limited to a small number of pioneering studies based on the high passage strains AD169 and Towne. I therefore sought to gain a more complete appreciation of the expression kinetics, size and post-translational modifications of the entire US12 family, as well as doing so in the more clinically relevant Merlin strain.

4.1 Analysis of US12 family members through the use of prediction software

Firstly the US12 family were analysed using prediction software to identify their potential traits. This included searching for N-glycosylation sites, organelle localisation motifs, and predicting their transmembrane domains and membrane topology. Most prediction software were created to identify motifs solely in eukaryotic or prokaryotic proteins, rather than viral ones. However, HCMV has evolved to replicate in human cells, so it would be natural for HCMV proteins to contain human motifs. Multiple software resources were used to provide an overview on the potential traits present in the US12 family.

4.1.1 N-glycosylation motifs of US12 family members

The prediction resource sites NetNGlyc (Gupta et al., 2004), Prosite (de Castro et al., 2006) and Eukaryotic Linear Motifs (ELM) (Dinkel et al., 2016) were used to analyse the US12 family members for the presence of N-glycosylation motifs. All 3 sites investigated the US12 family sequences for the N-glycosylation motif Asn-Xaa-Ser/Thr (Asparagine-Xaa-Serine/Threonine) or Asn-Xaa-Ser/Thr-Xaa where Xaa is any amino acid except for proline. ELM also searched for an additional atypical motif Asn-X-Cys (Cysteine).

It was universally predicted that US14, US16, US17 and US20 contained potential N-glycosylation sites (**Table 4.1**). US18 contained N-glycosylation motifs that were predicted not to be functional or not utilised for N-glycosylation within the cell due to either structural or conformational constraints. For example, the first site within US18 contained a proline just after the asparagine residue, which would cause conformational constraints, and its second site occurred within the transmembrane domain of the protein so would also not be glycosylated during infection (Dinkel et al., 2016). As one of US14's motif sites was scored below the threshold levels of the NetNGlyc software, it was also unlikely that this motif was N-glycosylated within the cell. US13 was predicted to have an atypical N-glycosylation motif that was predicted by ELM alone, and less is known about the

Table 4.1 N-glycosylation motifs predicted within the US12 family

Gene†	NetNGlyc*	Prosite*	ELM*
US12	-	-	-
US13	-	-	1 atypical motif: NIC (62-64aa).
US14	2 motifs: NGTL (291-294aa) and NSTT (299-302aa)✿	2 motifs: NGTL (291 - 294) and NSTT (299 - 302)	2 generic motifs: DNGTLS (290-295aa) and LNSTTA (298-303aa)
US15	-	-	-
US16	1 motif: NCTL (178-181aa) ✿	1 motif: NCTL (178-181aa)	1 generic motif: DNCTLS (177-181aa).
US17	1 motif: NLTR (287 – 290aa)	1 motif: NLTR: (287 – 290aa)	1 generic motif: RNLTRT (286-291aa)
US18	2 motifs: NPTR (159-162aa)✿, and NMSV (242-245aa) ✿	1 motif: NMSV (242-245aa)	-
US19	-	-	-
US20	3 motifs: NYSF (61 – 64aa), NATV (88 – 91aa) and NGTL (242 – 245aa)	3 motifs: NYSF (61-64), NATV (88-91), NGTL (242-245)	3 generic motifs: ENYSFF (60-65), SNATVL (87-92) and DNGTLT (241-246aa) and 2 atypical motifs: NFC, (56-58) and NQC (202-204aa)
US21	-	-	-

*NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/), PROSITE (ExPASy) resource portal (de Castro et al., 2006) and Eukaryotic Linear Motif (ELM) resource (Dinkel et al., 2016). ELM hits are those remaining after internal filtering by the software to remove motifs that are unlikely to be utilised in practice

† Protein sequences were translated from the DNA sequences of US12 family members from HCMV Merlin reference strain (NCBI RefSeq NC_006273.2).

✿ Unlikely to be utilised for N-glycosylation within the cell due to structural or conformational constraints

constraints and restrictions of this motif and the likelihood of this residue being glycosylated during infection. US12, US15, US19 and US21 were consistently predicted not to contain any N-glycosylation motifs.

4.1.2 Localisation motifs of US12 family members

Multiple prediction software, including Phobius (Kall et al., 2004), ELM (Dinkel et al., 2016) and Protter (Omasits et al., 2014) failed to detect any cleavable N-terminal signal peptides within US12 family members. However, the ELM resource did predict other motifs within the family that link to localisation, including non-cleavable ER retention motifs and lysosomal sorting signals. Although no US12 family were predicted to have the most common ER retention motif KDEL; US12, US13, US14, US15, US16, US17, US19 and US20 were all predicted to contain the ER retention motif TRG_ER_diArg_1 (**Table 4.2**). This motif was defined by two consecutive arginine (Arg/R) residues (RR) with or without a single residue insertion (RXR), and with an adjacent arginine or hydrophobic residue which may be on either side of the Arg pair. The functional motif needs to be exposed within a cytosolic region of the membrane protein and requires a distinct proximity to the transmembrane region, otherwise the motif is unused. Due to this, the ELM software filtered the results using globular domain filtering, structural filtering and context filtering to remove motifs that were unlikely to be used (Dinkel et al., 2016) and subsequently excluded US18 from the list of members likely to contain a functional ER retention motif. US21 was not predicted to contain any ER retention motifs by ELM, however an alternative motif was found within US21 using a different prediction site, Psort II (Nakai and Horton, 1999), which predicted an N terminus XXRR-like motif. Psort II also predicted US15 to contain an N-terminus CCRR-like motif and C terminus KKXX-like motif, US16 to have an N terminus XXRR-like motif, and US20 to contain a C terminus KKXX-like motif (**Table 4.2**). The XXRR-like motif is an N-terminal di-arginine motif, usually of type II proteins, and the KKXX-like motif is a C-terminal di-lysine motif, usually of type Ia proteins. Their existence is not sufficient for the localization of proteins to the ER membrane however, and the likelihood of these functioning as such is unknown.

Members US12, US13, US14, US18 and US20 were also predicted to contain the lysosomal targeting motif 'TRG_LysEnd_APsAcLL_1' with the consensus motif [Asp/Glu/Arg/Gln]-X-X-X-Leu-[Leu/Val/Ile], where X is any amino acid (**Table 4.2**). This is a sorting and internalisation signal directing transmembrane proteins (generally type I) from the cell surface or TGN to the lysosomal-endosomal compartment. The motif interacts with the

Table 4.2: Lysosomal targeting, ER retention and endocytic trafficking motifs predicted within the US12 family.

Gene†	ELM* hits	Psort II* hits
US12	4 x ER retention motif TRG_ER_diArg_1 and 1x lysosomal targeting motif TRG_LysEnd_APsAcLL_1	
US13	2x ER retention motif TRG_ER_diArg_1 and 2x lysosomal targeting motif TRG_LysEnd_APsAcLL_1	
US14	1x ER retention motif TRG_ER_diArg_1 and 1x lysosomal targeting motif TRG_LysEnd_APsAcLL_1	
US15	2x ER retention motif TRG_ER_diArg_1	1x XXRR-like ER retention motif and 1x KKXX-like ER retention motif
US16	5x ER retention motif TRG_ER_diArg_1	1x XXRR-like ER retention motif
US17	3x ER retention motif TRG_ER_diArg_1	
US18	2x lysosomal targeting motif TRG_LysEnd_APsAcLL_1	
US19	1x ER retention motif TRG_ER_diArg_1	
US20	2x ER retention motif TRG_ER_diArg_1 and 2x lysosomal targeting motif TRG_LysEnd_APsAcLL_1	KKXX-like ER retention motif
US21	None	XXRR-like ER retention motif

*Eukaryotic Linear Motif (ELM) resource (Dinkel et al., 2016) and Psort II programs (Nakai and Horton, 1999). ELM hits are those remaining after internal filtering by the software to remove motifs that are unlikely to be utilised in practice.

† Protein sequences were translated from the DNA sequences of US12 family members from HCMV Merlin reference strain (NCBI RefSeq NC_006273.2).

adaptor proteins AP1, AP2, AP3 and probably AP4, although it is unclear how the specificity for binding to APs is achieved. US16, US19 and US21 also contained the sequence for this motif, but were removed from the prediction after ELM filtering as previously described (Dinkel et al., 2016).

Some members also contained the sequence for the TRG_ENDOCYTIC_2 motif which is a tyrosine (Y)- based sorting signal that can direct trafficking within the endosomal and the secretory pathways, however none of these remained as motifs that were predicted to be utilised after ELM filtering had taken place (Dinkel et al., 2016), except for US20 (**Table 4.2**).

4.1.3 Transmembrane domain and topology predictions of US12 family members

The US12 family are predicted to each encode a polypeptide that passes through a lipid membrane multiple times. Multiple prediction websites predicted between 6 and 8 transmembrane (TM) domains for the US12 family members (**Table 4.3**). Often it was the presence of a less hydrophobic region, the possibility of a re-entrant loop instead of a TM domain or conflicting membrane orientations that caused the differences in the predictions between resources. Nevertheless, the consensus across prediction sites was that each member had 7 TM domains (**Table 4.3**). In line with these predictions, the protein topology prediction software Protter (Omasits, Ahrens, Müller, & Wollscheid, 2014) also predicted 7TM domains for each family member, and provided a unique non-linear perspective on how the US12 family members might cross the membrane (**Figure 4.1**). The mid-section of each of the proteins (between TM domains 1 and 7) was observed to contain mostly intra-membrane regions with only small looped external sections between each TM domain. Generally, the longest regions of each protein that were not within the membrane appeared to be at the N and C-termini, except for US19 which had a very short exposed C-terminus (**Figure 4.1**). Predicted N-glycosylation sites were also indicated. The Protter N-termini orientation of each US12 family member was manually selected based on the consensus across all prediction sites (**Table 4.3**). These included Psort (Nakai and Horton, 1999), TMPred (Hofmann and Stoffel, 1993) and the Constrained Consensus Topology prediction server (CCTOP) (Dobson et al., 2015) which compared 10 methods of membrane topology prediction in order to reach a consensus. It was revealed that the prevalent prediction was that the N-terminus of the US12 family members would be cytoplasmic, and that the C-terminus would be non-cytoplasmic (**Table 4.3**). N- and C- termini orientations are generally predicted using either the 'positive-inside' rule where cytoslic loops near the lipid bilayer are known to contain more positively charged amino acids or by the by the net charge difference of the 15 residues flanking the most N-terminal transmembrane

Table 4.3: US12 family predictions of transmembrane domains and membrane topology

	CCTOP* (consensus from 10 sites)	TMpred*	Psort*
US12	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	7TMDs, with the N-terminus inside (cytoplasmic) and C-terminus outside	6 TMDs, with N-terminus outside (non- cytoplasmic)
US13	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	7TMDs, with the N-terminus inside (cytoplasmic) and C-terminus outside	7TMDs, with N-terminus inside (cytoplasmic)
US14	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	7TMDs, with the N-terminus inside (cytoplasmic) and C-terminus outside	6TMDs, with N-terminus inside (cytoplasmic)
US15	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	6TMDs , with the N-terminus inside (cytoplasmic) and C-terminus inside	7TMDs, with N-terminus inside (cytoplasmic)
US16	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	7TMDs, with the N-terminus inside (cytoplasmic) and C-terminus outside	8TMDs, with N-terminus outside (non-cytoplasmic)
US17	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	7TMDs, with the N-terminus inside (cytoplasmic) and C-terminus inside	7TMDs, with N-terminus inside (cytoplasmic)
US18	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	7TMDs, with the N-terminus outside (non-cytoplasmic) and C-terminus inside	7TMDs, with N-terminus inside (cytoplasmic)
US19	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	7TMDs, with the N-terminus outside (non-cytoplasmic) and C-terminus inside	6TMDs, with N-terminus inside (cytoplasmic)
US20	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	7TMDs , with the N-terminus inside (cytoplasmic) and C-terminus outside	7TMDs, with N-terminus inside (cytoplasmic)
US21	7TMDs with N-terminus cytoplasmic and C-terminus non-cytoplasmic	8TMDs, with the N-terminus outside (non-cytoplasmic) and C-terminus outside	7TMDs, with N-terminus inside (cytoplasmic)

 * The Constrained Consensus Topology CCTOP prediction server (Dobson et al., 2015) compared 10 methods of membrane topology prediction in order to reach a consensus. The TMpred server (Hofmann and Stoffel, 1993) classified 'inside' as normally meaning the cytoplasmic face, and 'outside' as the lumenal face of the membrane, depending on the organelle. Psort (Nakai and Horton, 1999) predicted that the more positive portion of the protein faces the cytosol.

Figure 4.1

US16

US17

US18

US19

Figure 4.1 cont.

US20 **US21**
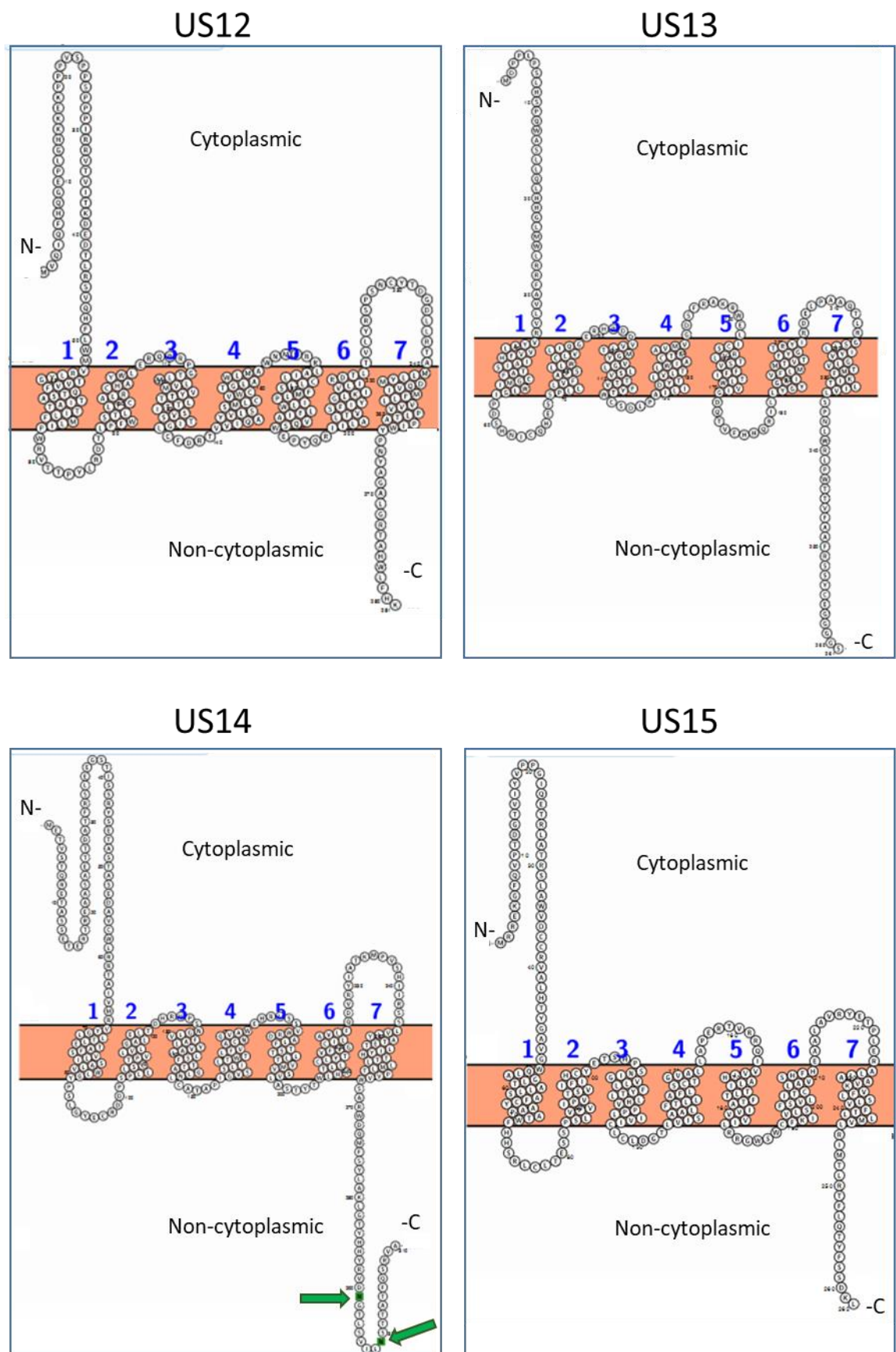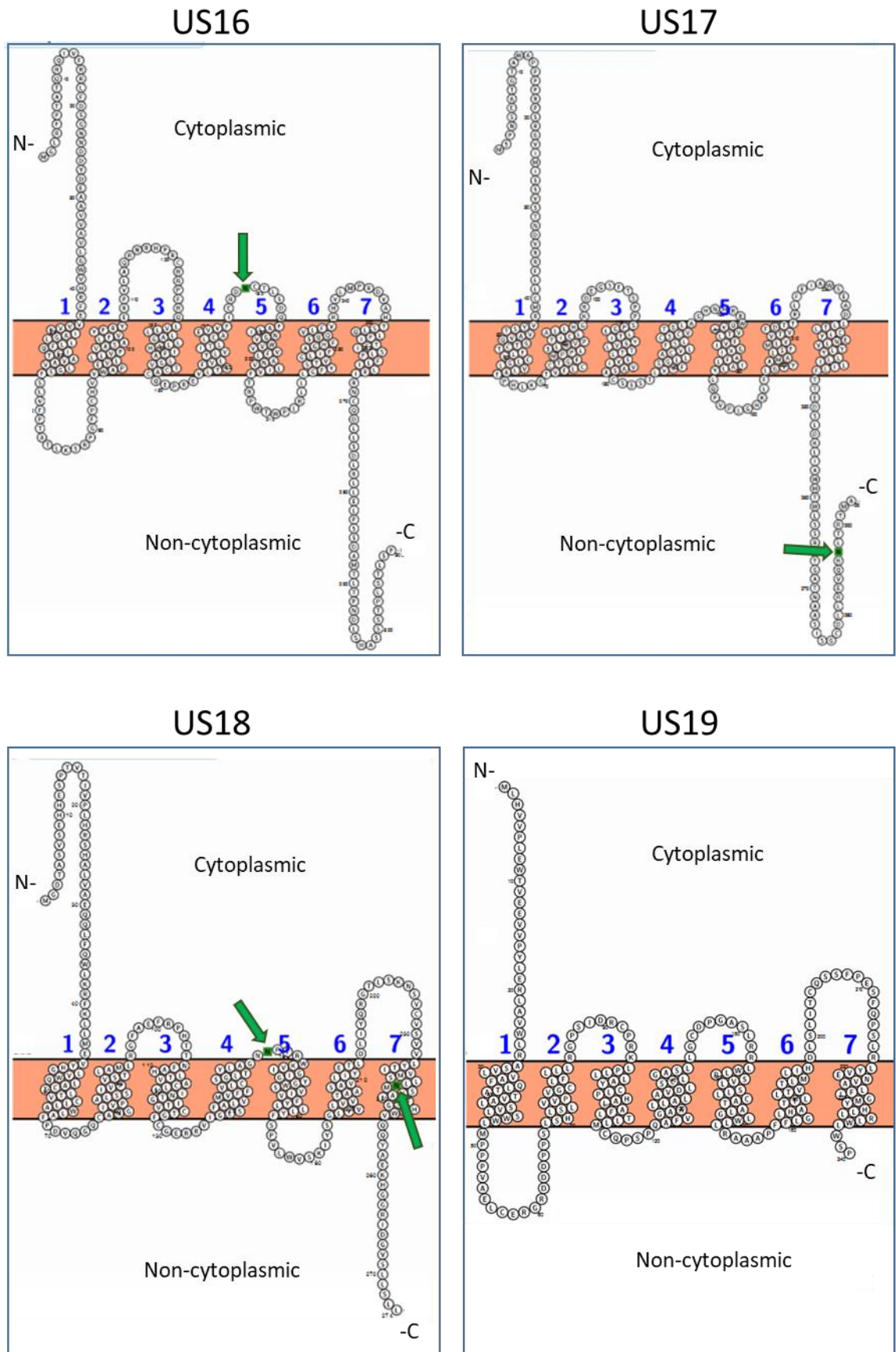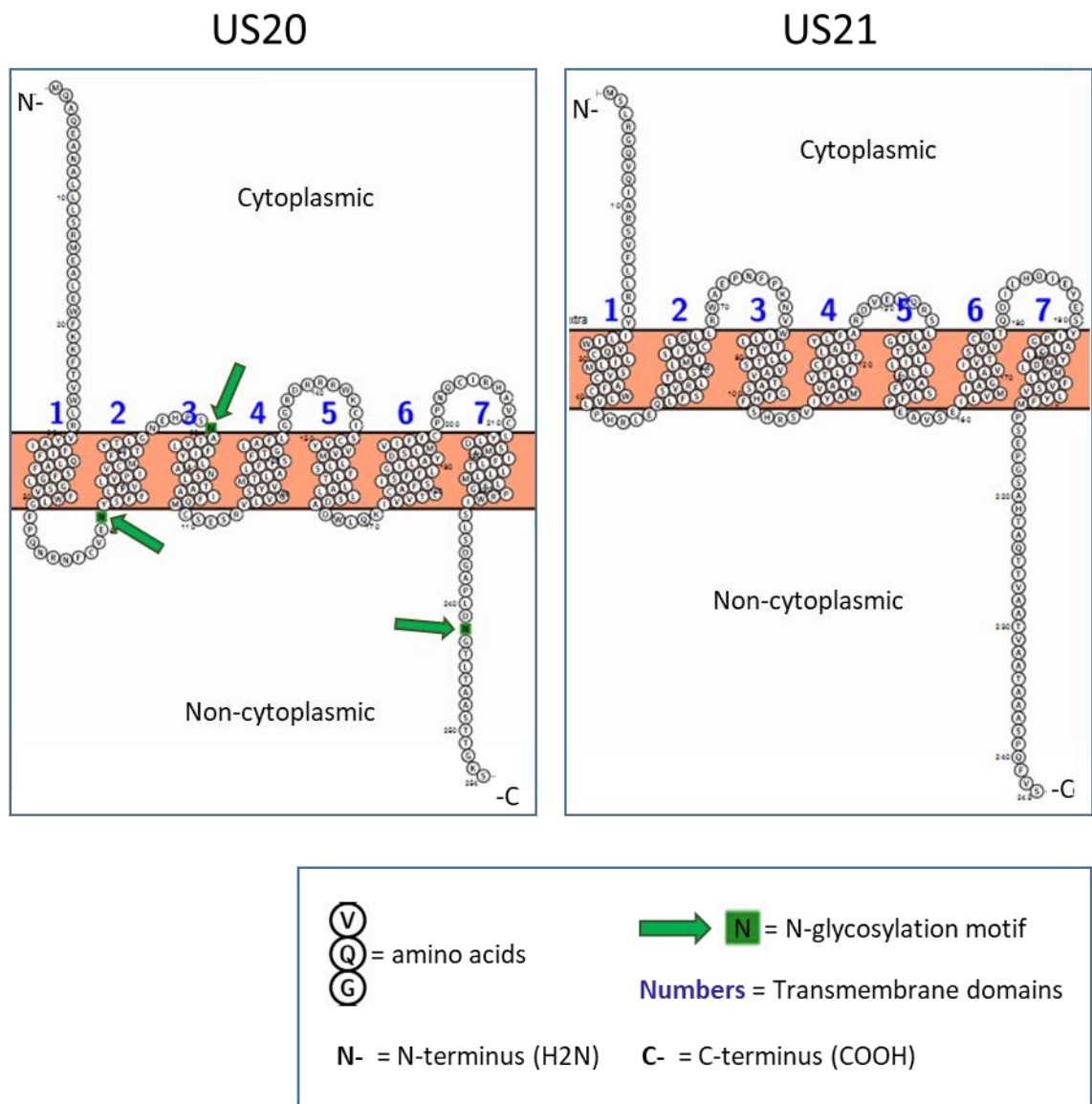
Figure 4.1: Predicted topology structures of the US12 family members. US12 family protein topologies were determined from the amino acid sequence of each member when inputted into Protter software (Omasits, Ahrens, Müller, & Wollscheid, 2014) which allowed the predicted topology and featured annotations to be visualised. For each member, the software allows you to visualise the predicted 7-transmembrane domains (numbered 1-7 in purple), along with any N-glycosylation sites predicted (in green, with arrows) in a non-linear manner. N-termini were manually orientated to correspond with the consensus that the N-terminus of each member was most likely to be cytoplasmic (**Section 5.1.3**)

segment. Psort predicted that the more positive portion of the protein would face the cytosol, and the TMpred server classified the 'inside' orientation as normally meaning the cytoplasmic face, and 'outside' as the lumenal face of the membrane, depending on the organelle. 'Outside' or 'non-cytoplasmic' may also refer to the termini being extracellular facing if the protein was found to localise at the plasma membrane.

4.2 Kinetics of HCMV US12 family member expression

Having established that all V5-tagged US12 family members could be detected by immunoblot (**Chapter 3, Fig. 3.6**), their expression kinetics were studied at 4 time-points over the course of a productive HCMV infection (**Fig. 4.2**). The immunoblot expression kinetics of US12 family members were also compared to the average temporal class patterns from Weekes et al. (2014), as demonstrated by **Figure 1.5 (Section 1.4.1.4)**. Where data was available, a comparison was also made with the 'quantitative temporal viromics' (QTV) evaluation of the US12 family members' temporal class and presented graphs were generated by interrogating the QTV data from supplemental Table S2 (Weekes et al., 2014). Densitometry calculations were also made from the immunoblot data and these output data were plotted as line graphs of relative expression over time (**Fig. 4.3**), in order to make these comparisons easier.

Immunoblotting detected all ten US12 family members compared to a combined total of only six US12 proteins by QTV across both whole cell lysate (WCL) and plasma membrane (PM) samples (**Fig. 4.2**). The ability of the QTV study to detect pUS15, pUS17 and pUS18 varied across experiments, with these 'partially detected' proteins lacking detection in either the WCL or the PM QTV samples.

pUS13-V5, pUS18-V5 and pUS19-V5 were designated as being Tp3 proteins on the basis of the immunoblot data, with all three proteins readily detected at 48 hpi and their expression increasing over time and only levelling out at 96 hpi (**Fig. 4.2**), thus following the average pattern of Tp3 proteins. pUS21-V5 was assigned to Tp5 as it was barely detectable at 48 hpi by immunoblot with its expression subsequently increasing over time right up until 96 hpi (**Fig. 4.2**). To further categorise pUS21-V5, its expression pattern could be observed with the addition of the viral replication inhibitor phosphonoformate (PFA), as late proteins, including Tp5 proteins, should be majorly reduced in its presence.

US12-V5

QTV data from
Weekes et al., 2014

V5

Actin

IE1

UL44

UL99

24 48 72 96

QTV
WCL

QTV
WCL
+/-PFA

QTV
PM

US13-V5

V5

Not detected

QTV
WCL

Actin

Not detected

QTV
WCL
+/-PFA

IE1

UL44

Not detected

QTV
PM

UL99

Fig. 4.2   24 48 72 96

Key:
No PFA   WCL
With PFA   PM

US14-V5

QTV WCL

QTV WCL +/-PFA

QTV PM

V5

Actin

IE1

UL44

UL99

24 48 72 96

US15-V5

QTV WCL

QTV WCL +/-PFA

Not detected

QTV PM

V5

Actin

IE1

UL44

UL99

24 48 72 96

Fig. 4.2 cont.

Key:
- No PFA
- With PFA
- WCL
- PM

US16-V5

60 —				
40 —				V5
30 —				
20 —				

Not detected — **QTV WCL**

40 —	Actin

Not detected — **QTV WCL +/-PFA**

60 —	IE1

40 —	UL44

Not detected — **QTV PM**

30 —	UL99

24 48 72 96

QTV data from Weekes et al., 2014

US17-V5

60 —				
40 —				V5
30 —				
20 —				

QTV WCL

(graph: y-axis 0 to 1, x-axis 0, 24, 48, 72, 96)

40 —	Actin

Not detected — **QTV WCL +/-PFA**

60 —	IE1

40 —	UL44
30 —	

Not detected — **QTV PM**

30 —	UL99
20 —	

24 48 72 96

Fig. 4.2 cont.

Key:

—●— No PFA		—●— WCL	
—●— With PFA		—●— PM	

US18-V5

60 —	
40 —	
30 —	
	V5
20 —	
15 —	

Not detected — **QTV WCL**

40 —	**Actin**

60 —	**IE1**

QTV WCL +/-PFA

40 —	
30 —	**UL44**

Not detected — **QTV PM**

30 —	
20 —	**UL99**

24 48 72 96

US19-V5

60 —	
40 —	
30 —	
20 —	**V5**
15 —	

Not detected — **QTV WCL**

40 —	**Actin**

Not detected — **QTV WCL +/-PFA**

60 —	**IE1**

40 —	
30 —	**UL44**

Not detected — **QTV PM**

30 —	
20 —	**UL99**

Fig. 4.2 cont.   24 48 72 96

Key:
- No PFA
- With PFA
- WCL
- PM

US20-V5

QTV data from Weekes et al., 2014

V5

Actin

IE1

UL44

UL99

QTV WCL

QTV WCL +/-PFA

QTV PM

24 48 72 96

US21-V5

V5

Actin

IE1

UL44

UL99

Not detected — QTV WCL

Not detected — QTV WCL +/-PFA

Not detected — QTV PM

Key:
No PFA — WCL
With PFA — PM

Fig. 4.2 cont.   24 48 72 96

Figure 4.2: **Immunoblot time-course of US12 family protein expression compared to the US12 family data from the quantitative temporal viromics (QTV) study.** Fibroblasts were mock infected or infected with HCMV Merlin, or HCMVs encoding V5 tagged US12 genes at an MOI of 10. Samples were left untreated and harvested at 24, 48, 72 and 96 hpi. Whole cell lysates were extracted and proteins detected by immunoblotting using an anti-V5 antibody and an HRP secondary antibody. All samples were also tested with anti-actin used as a loading control, IE1 as a Tp1 protein control, and UL44 and UL99 as Tp5 protein controls. Immunoblots have a side-by-side comparison with supplementary QTV proteomics data collected on the US12 family found in Table S2 (Weekes et al., 2014). Fibroblasts were mock infected or infected with HCMV Merlin at a MOI of 10. For each member, whole cell lysate (WCL, encoded in orange) and plasma membrane (PM, coded in blue) samples were collected as described in (Weekes et al., 2014). Line graphs were represented as relative expression (0-1) over time. For one experiment, phosphonoformate (PFA) was added at 300ug/ml from the time of infection onward (encoded in purple). Proteins were identified by mass spectrometry.

US12-V5

US17-V5

US13-V5

US18-V5

US14-V5

US19-V5

US15-V5

US20-V5

US16-V5

US21-V5

Relative expression by densitometry

Time (h.p.i)
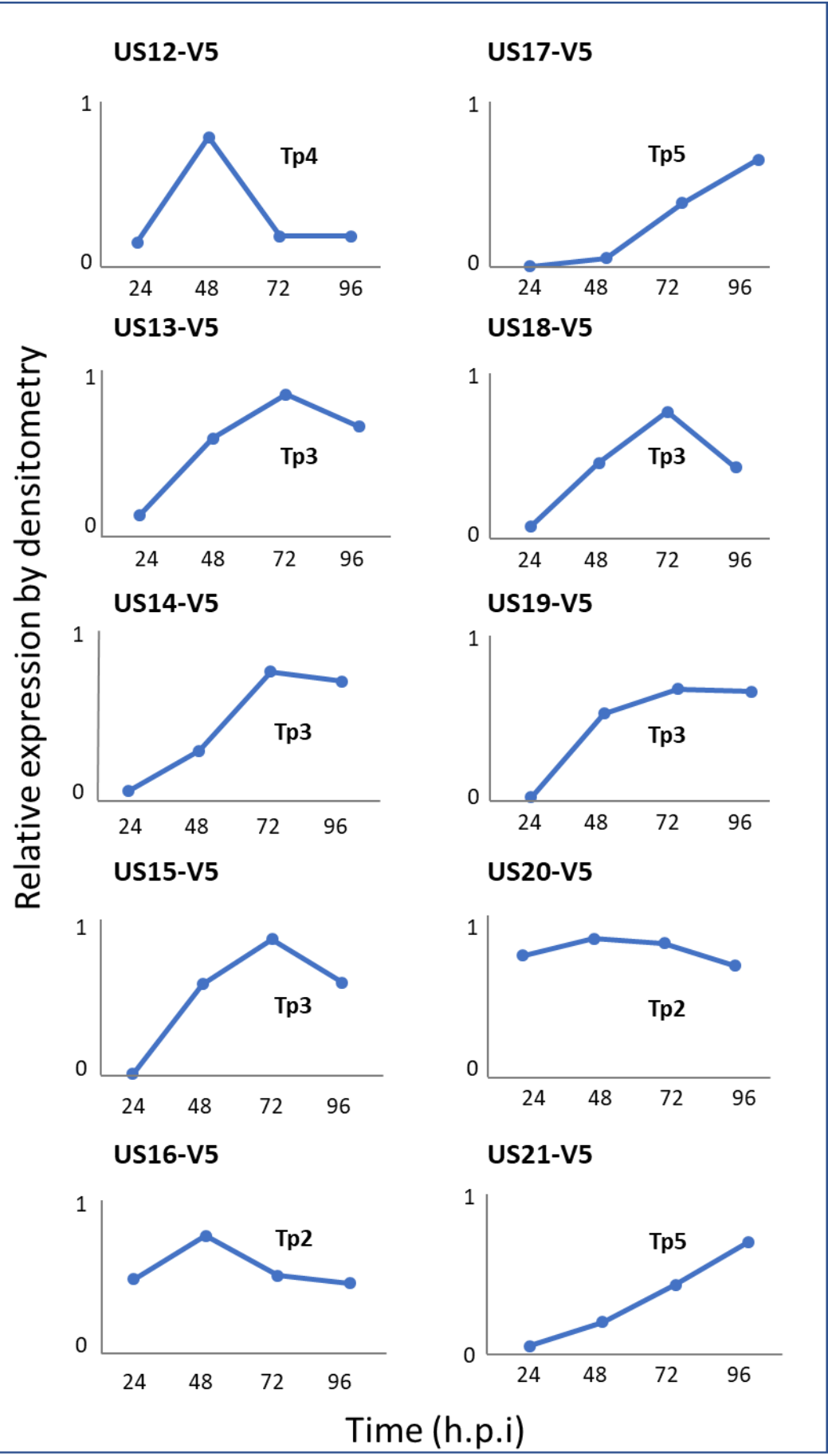
Figure 4.3: Densitometry time-course of US12 family protein expression from immunoblot data in Figure 4.2. In order to graphically represent the US12 family total protein expression over time, the GeneTools program (www.syngene.com/software/genetools-automatic-image-analysis/) was used to calculate the densitometry of each US12 family member at each time-point (24, 48, 72 and 96 hpi) from the whole cell lysate (WCL) immunoblot time-course data (**Figure 4.2**). The densitometry output data for each US12 family member were then plotted as relative expression over time. This allowed for an easier comparison between the data in this thesis and the expression of each US12 family member from the proteomics quantitative temporal viromics (QTV) study (Table S2 in Weekes et al. (2014) as depicted in **Figure 4.3**). US12 family members were assigned a temporal kinetics class (Tp1-5) through the comparison of their expression over time and the average temporal profiles for proteins in each temporal class (**Figure 1.5**).

pUS16-V5 and pUS20-V5 (**Fig. 4.2**) both had high levels of protein expression from early time-points (24 hpi) and generally stayed at high levels across the immunoblot time-course, following the average pattern of Tp2 proteins. This designation supported the categorisation of pUS20 as a Tp2 protein by QTV. Some late proteins can be detected at early time-points due to their presence in input virions, as frequently demonstrated by the early (24 hpi) detection of the Tp5 structural protein pp28/ppUL99 (**Fig. 4.2**). However, the pp28 levels for pUS20-V5, and especially for pUS16-V5, are relatively low at 24 hpi so their presence was likely to be at least partially due to *de novo* protein expression.

pUS12 was also previously designated by QTV to be a Tp2 protein, and its immunoblot expression pattern did show some similarities to the average Tp2 protein profile. However, it demonstrated more similarities to the distinct Tp4 pattern, with maximal expression at 48 hpi and low expression at other time-points so has been classified as a Tp4 protein by immunoblot (**Fig. 4.2**). Observing pUS12 expression with and without PFA would further substantiate this as viral DNA inhibition should have little effect on early proteins, but will partially inhibit the majority of late proteins.

pUS17-V5 was expressed especially late in infection, with very low expression levels at 48 hpi which subsequently increased over time, peaking at 96 hpi (**Fig. 4.2**). This was comparable to the average pattern of Tp5 proteins so pUS17 was categorised as such. pUS17 was also categorised as a late protein by QTV but was instead designated as a Tp4 protein as its peak expression by QTV was earlier at 48 hpi.

Immunoblotting detected pUS14-V5 at low levels from 24 hpi, increasing steadily over time, so it was accordingly categorised as a Tp3 protein (**Fig. 4.2**). This correlated with its designation as a Tp3 protein by QTV. pUS15-V5 was also suggested to be a Tp3 protein through immunoblot analysis, as its protein expression was detected from 48 hpi and seen to increase over time, peaking at 72 hpi. However pUS15 was instead designated as a Tp2 protein by QTV as it had demonstrated an extra peak in expression prior to 24 hpi. Collecting and immunoblotting earlier time-points would further determine whether pUS15-V5 was in fact a Tp2 or Tp3 protein.

The immunoblot data presented here generally agreed with the QTV data, with both datasets providing the same classifications for pUS14 and pUS20, and similar patterns of expression seen over time for pUS12, pUS15 and pUS17 (**Table 4.4**). The other family members had not previously been assigned a temporal class so these classifications were novel. The inclusion of additional time-points to provide more coverage earlier in infection

Table 4.4: Summary of the comparison between the US12 family data from the immunoblot time-course and the QTV proteomics study, alongside previously published results.

Gene	Time-course expression pattern*	QTV expression pattern*	Previously published immunoblot results
US12	Tp4 pattern	Tp2 protein	N/A†
US13	Tp3 pattern	Not detected	N/A†
US14	Tp3 pattern	Tp3 protein	N/A†
US15	Tp3 pattern	Tp2 protein	N/A†
US16	Tp2 pattern	Not detected	Detected at 48, 72 and 96 hpi (Luganini et al., 2017, Bronzini et al., 2012).
US17	Tp5 pattern	Tp4 protein	
US18	Tp3 pattern	Not assigned a temporal class	Detected from 34 hpi (Guo, 1993).
US19	Tp3 pattern	Not detected	Detected at a low level at 18 hpi and detected until 72 hpi (Guo, 1993).
US20	Tp2 pattern	Tp2 protein.	Detected from 18 hpi and until 72 hpi (Guo, 1993), and detected equally from 24 to 96 hpi (Cavaletto et al., 2015).
US21	Tp5 pattern	Not detected	N/A†

* In relation to **Figure 4.2**. QTV data can also be found in Table S2 by Weekes et al. (2014).

† N/A, not applicable as no previously published results

and the utilisation of an inhibitor of virus replication (e.g. PFA) would further substantiate the expression kinetics of US12 family members.

All US12 family member proteins could be detected by 48 hpi, while by 72 hpi most were near peak levels of expression. By 72 hpi, strain Merlin has also established a full productive infection in fibroblasts, with viral DNA replication activated, late genes being expressed and viral immune evasion functions being deployed. For these reasons, the 72h time-point was favoured in later experiments.

Although a similar pattern of HCMV IE1, UL44 and UL99 gene expression appears relatively similar across all V5-tagged viruses and the HCMV 1111 control virus, it cannot be formally ruled out that the V5 tag is not having an effect on the growth of the viruses without first undertaking growth curve experiments.

4.3 N-glycosylation states of the US12 family members when expressed in isolation

To simplify the analysis, I sought to investigate whether any of the US12 family proteins were subject to N-glycosylation using a set of replication-deficient adenovirus (Ad) recombinants encoding the ten US12 family members, all with a C-terminal V5 tag (produced by Dr S. Seirafian). The fibroblasts used in this study were engineered to overexpress the Coxsackie-adenovirus receptor (HF-CARs) to facilitate Ad vector delivery. Cell extracts were then treated with EndoH or PNGase F to remove N-glycans before being analysed by SDS-PAGE and immunoblot to detect any alteration in mobility of the protein resulting from glycosidase treatment.

The majority of US12 family members expressed from Ad vectors had predominant protein species, yet multiple extra protein forms and/or smear-type patterns were also seen (**Fig. 4.4**). The only infection for which multiple protein bands or smears could not be detected was with Ad US21-V5, which solely presented with its 2 main protein species of 17 and 17.5 kDa. The apparent molecular weights of the US12 family are presented later in **Table 4.5 (Section 4.5)**. Ad pUS15-V5 and pUS18-V5 were not detected in the untreated or EndoH treated samples, so it was impossible to compare their molecular weights between treatment types to determine whether they were N-glycosylated. All other US12 family members expressed from Ad vectors showed no reduction in molecular weight of their main protein form(s) and were therefore non-N-glycosylated proteins (**Fig. 4.4**). The only

RAd US12-V5   **RAd US13-V5**   **RAd US14-V5***

| Key: |
| U= Untreated |
| E= EndoH |
| P= PNGase F |

V5

Actin

U E P   U E P   U E P

RAd US15-V5   **RAdUS16-V5**   **RAd US17-V5***

V5

Actin

U E P   U E P   U E P

Fig. 4.4

RAd US18-V5   RAd US19-V5   RAd US20-V5

Key:
U= Untreated
E= EndoH
P= PNGase F

V5

Actin

U   E   P   U   E   P   U   E   P

RAd US21-V5   RAd2153   Mock

V5

Actin

152

U   E   P   U   E   P   U   E   P   Fig. 4.4 cont.
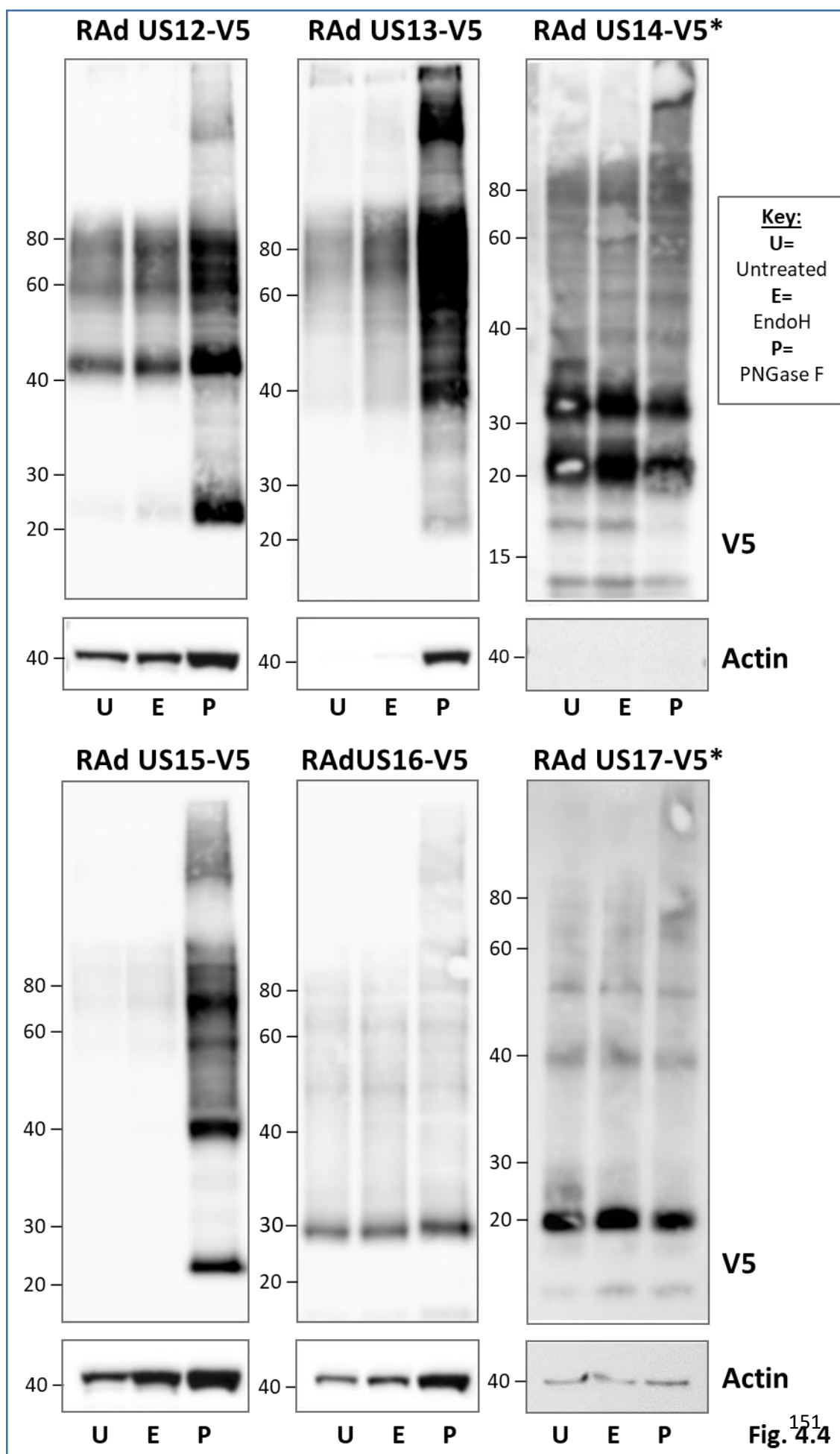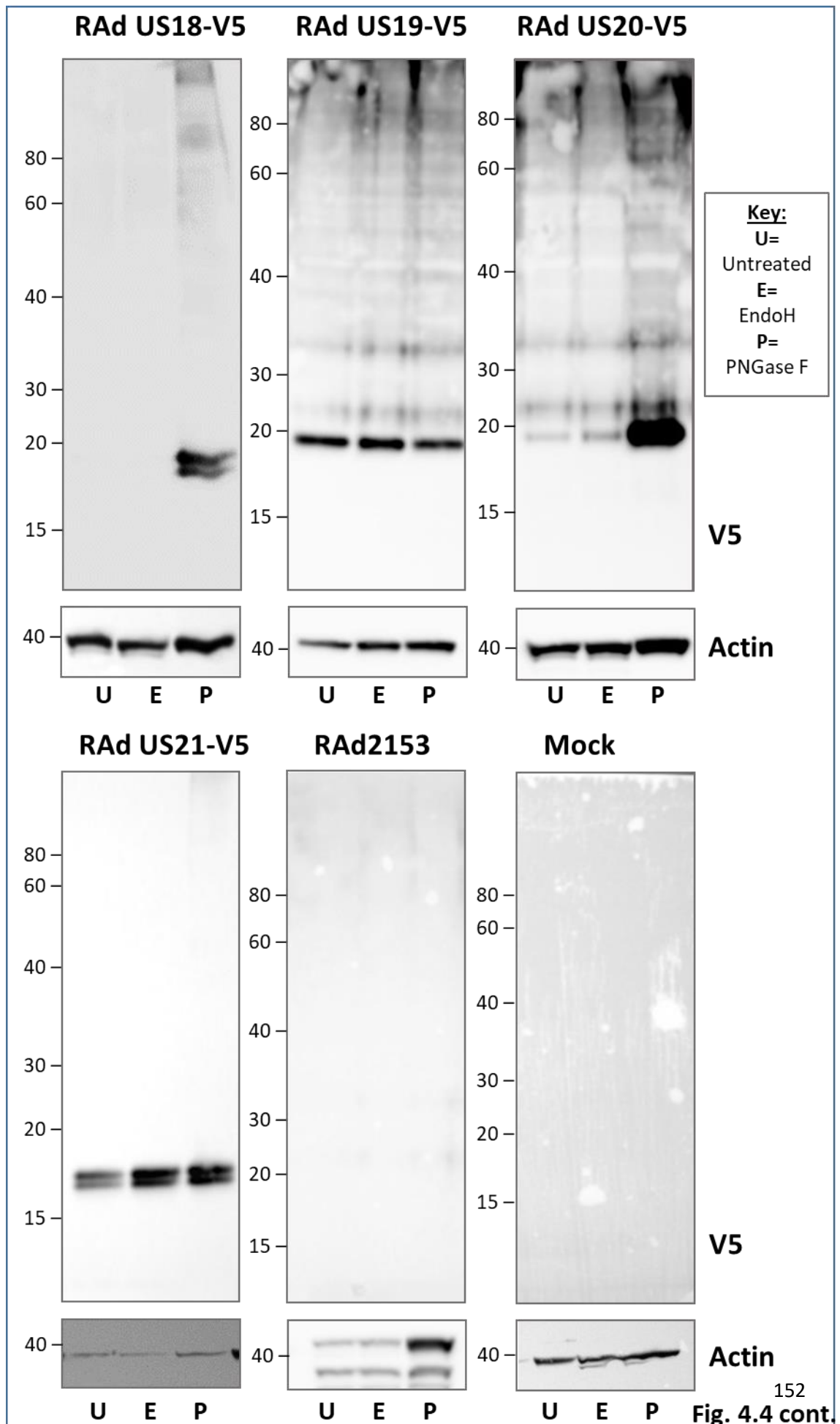
Figure 4.4: The N-glycosylation patterns of V5-tagged US12 family members as expressed from an adenovirus expression system. Fibroblasts (HF-CARS) were mock infected or infected with replication-deficient adenovirus (Rad) 1253, or RAds encoding V5 tagged US12 genes an MOI of 10. Samples were harvested at 72h.p.i. Proteins were extracted using the whole cell lysate method. Samples were either left untreated, or treated with EndoH or PNGase F treatment overnight. Samples were immunoblotted and detected using an anti-V5 antibody and an HRP secondary antibody. All samples were run with anti-actin as a loading control. All samples are representative of 2 independent experiments. *US14-V5 and US17-V5 RAds had been codon optimized and therefore had much higher expression levels

differences in molecular weight seen between the treatments was observed in the minor protein species of Ad pUS14-V5 and pUS17-V5 at 72 hpi, which do not represent the protein as a whole. pUS14-V5 appeared to have 2 minor species which displayed N-glycosylation, with a ~37 kDa EndoH sensitive glycoform and a ~17 kDa EndoH resistant form (**Fig. 4.4**). This implied that the 17 kDa protein glycoform had been processed through the Golgi, and that the ~37 kDa glycoform had yet to pass through the ER. pUS17-V5 similarly appeared to have an N-glycosylated low abundance protein species of 24 kDa that was sensitive to both EndoH and PNGase F (**Fig. 4.4**). Both US14-V5 and US17-V5 had previously been codon optimised within the Ad vector, which may have affected their expression and modifications within this system. This also caused them to have much higher expression levels than the other US12 family members so these N-glycosylated protein species may have not been detected or present under non-optimised conditions.

There also remains the possibility that Ad pUS12-V5 could be partially N-glycosylated, as its lower ~22 kDa protein form could be scarcely detected in untreated or EndoH conditions, but appears to be highly abundant after PNGase F treatment. However, as US12 is not predicted to contain any N-glycosylation motifs (**Table 4.1**), this is improbable. Instead, this difference is more likely caused by differential detection due to differential loading or a difference in treatment conditions. A similar effect was also seen with Ads pUS13-V5, pUS15-V5, pUS18-V5 and pUS20-V5 (**Fig. 4.4**). Alongside US12, US15, US19 and US21 also did not contain any N-glycosylation motifs (**Table 4.1**).

Overall, there is minimal evidence for N-glycosylation of these proteins in isolation which was unexpected as there are predicted N-glycosylation sites within some US12 family members**.** US13, US16, US18 and US20 were all predicted to contain N-glycosylation motifs, yet none of these proteins showed signs of N-glycosylation (**Table 4.1, Fig. 4.4**). US14 and US17 were also predicted to contain N-glycosylation motifs so it was surprising that only minor subsets appeared to be N-glycosylated.

It may be that these N-glycosylation sites are not used, or it may be that expressing the proteins in isolation does not lead to the correct processing of these US12 family members. Using the adenovirus as an expression system is advantageous due to the ease of recombineering and of growing the viruses, but it doesn't seem to provide reliable results for these transmembrane proteins without the context of infection, which is highlighted by the differences seen in comparison to the HCMV time-course (**Fig. 4.2**).

4.4 N-glycosylation states of the US12 family members in the context of HCMV infection

To follow up the N-glycosylation results of the US12 family proteins in isolation, I wanted to assess whether the US12 family member proteins were N-glycosylated in the context of HCMV infection. HCMV Merlin viruses containing the V5-tagged US12 members were used to infect fibroblasts (HF-TERTs), with WCLs harvested at 72 hpi. These were either left untreated or were treated with EndoH or PNGase F, then detected with an anti-V5 antibody. HCMV UL141 was known to be N-glycosylated and was ER-localised and therefore sensitive to both enzymes so was used as a positive control (Tomasec et al., 2005).

Seven out of 10 of the US12 family members appeared not to be N-glycosylated in the context of HCMV infection, with only pUS14-V5, pUS17-V5 and pUS20-V5 demonstrating N-glycosylation (**Fig. 4.5**). HCMV pUS14-V5 was fully N-glycosylated, with both EndoH and PNGase F able to remove the N-glycosylation, causing a reduction in size from ~37 kDa to ~33 kDa. HCMV pUS17-V5 however was only partially glycosylated, with only the predominant protein species of pUS17-V5 having N-glycosylation which could also be removed by both EndoH and PNGase F. This reduced the predominant higher molecular weight protein species (~26 kDa) to the size of the lower molecular weight species (23 kDa) (**Fig. 4.5**). Similarly, HCMV pUS20-V5 was partially N-glycosylated, although in this case it was the minor subset of the protein that was N-glycosylated and this could again be removed by both EndoH and PNGase F. This reduced the higher molecular weight protein (~23 kDa) of pUS20-V5 to ~19 kDa. All 3 of these proteins were predicted to contain functional N-glycosylation sites (**Table 4.1**). It was also predicted by structural models that these sites would be exposed to the lumen of the ER where they would be exposed to the glycosylation 'machinery' (**Fig. 4.1**). The data presented here therefore demonstrated that 3 members of the US12 family were N-glycosylated at 72 hpi when expressed in the context of HCMV infection. The fact that the glycosylation was EndoH-sensitive inferred that these proteins had immature glycosylation, thus had not transited through the Golgi apparatus.

All other US12 family members displayed no change in molecular weight (**Fig. 4.5**), and these proteins therefore appeared un-glycosylated. Some of these members did appear to have potential N-glycosylation sites however (**Table 4.1)** although not all of these sites were predicted to be utilised or functional, and US13 contained an atypical N-glycosylation site only (**Section 4.1.1**). Therefore, US16 remained the only member that was predicted to

US12-V5 **US13-V5** **US14-V5** **US15-V5**

US16-V5 **US17-V5** **US18-V5** **US19-V5**

Key: **U**= Untreated; **E**=EndoH; **P**=PNGase F

Figure 4.5

US20-V5   **US21-V5**   **HCMV**   **Mock**

V5

Actin

UL141

U E P   U E P   U E P   U E P
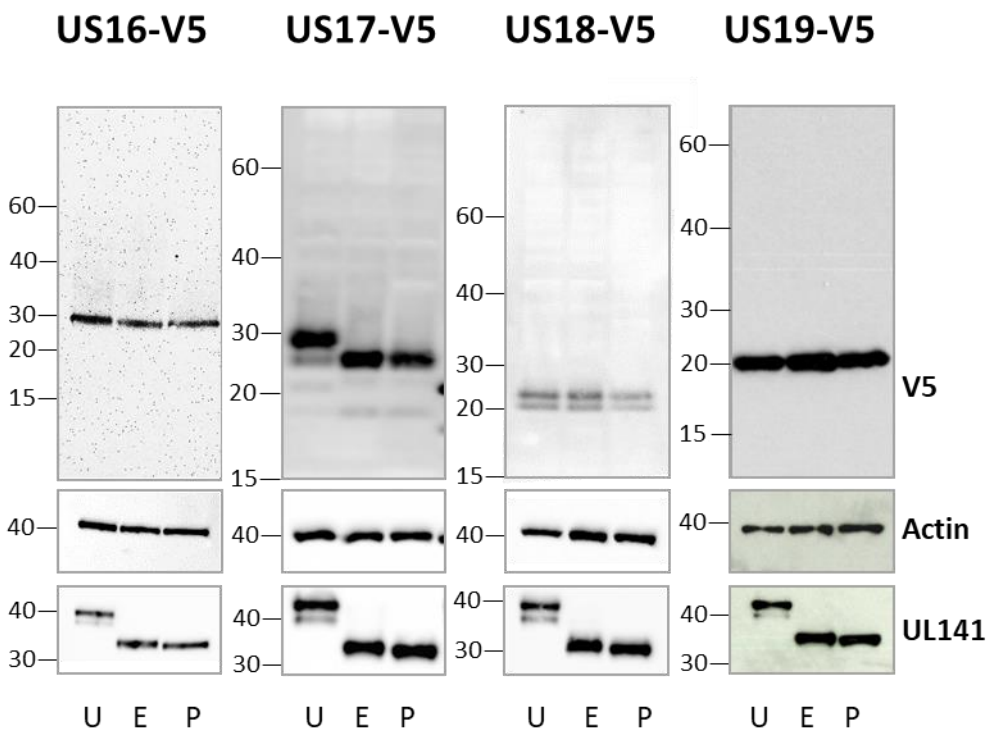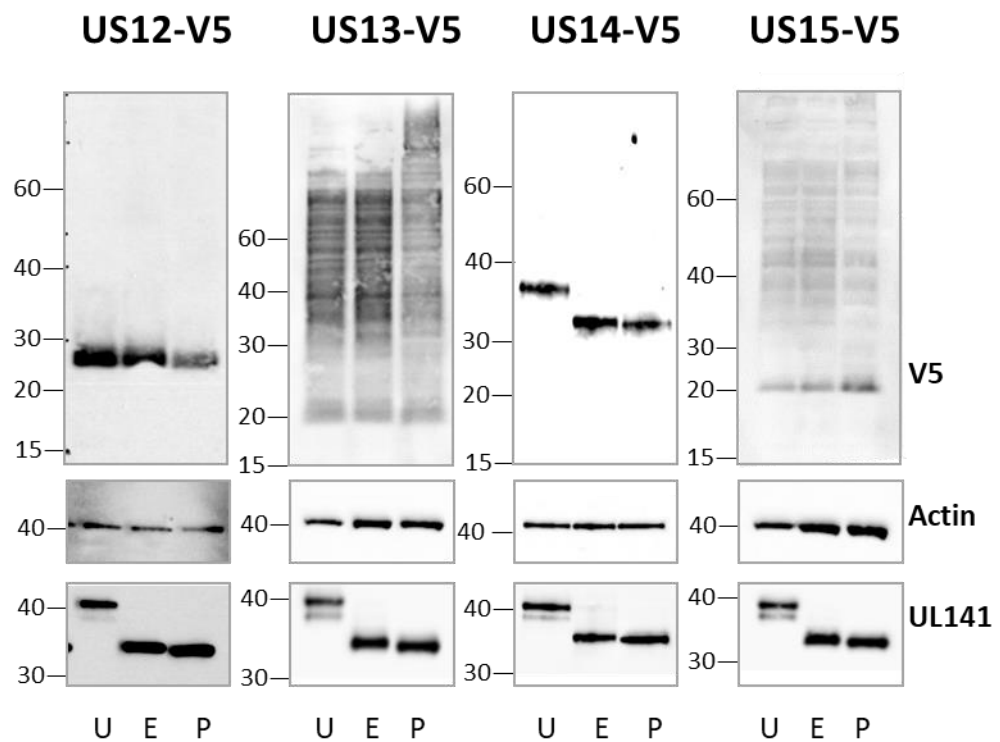
Key: U= Untreated; **E**=EndoH; **P**=PNGase F
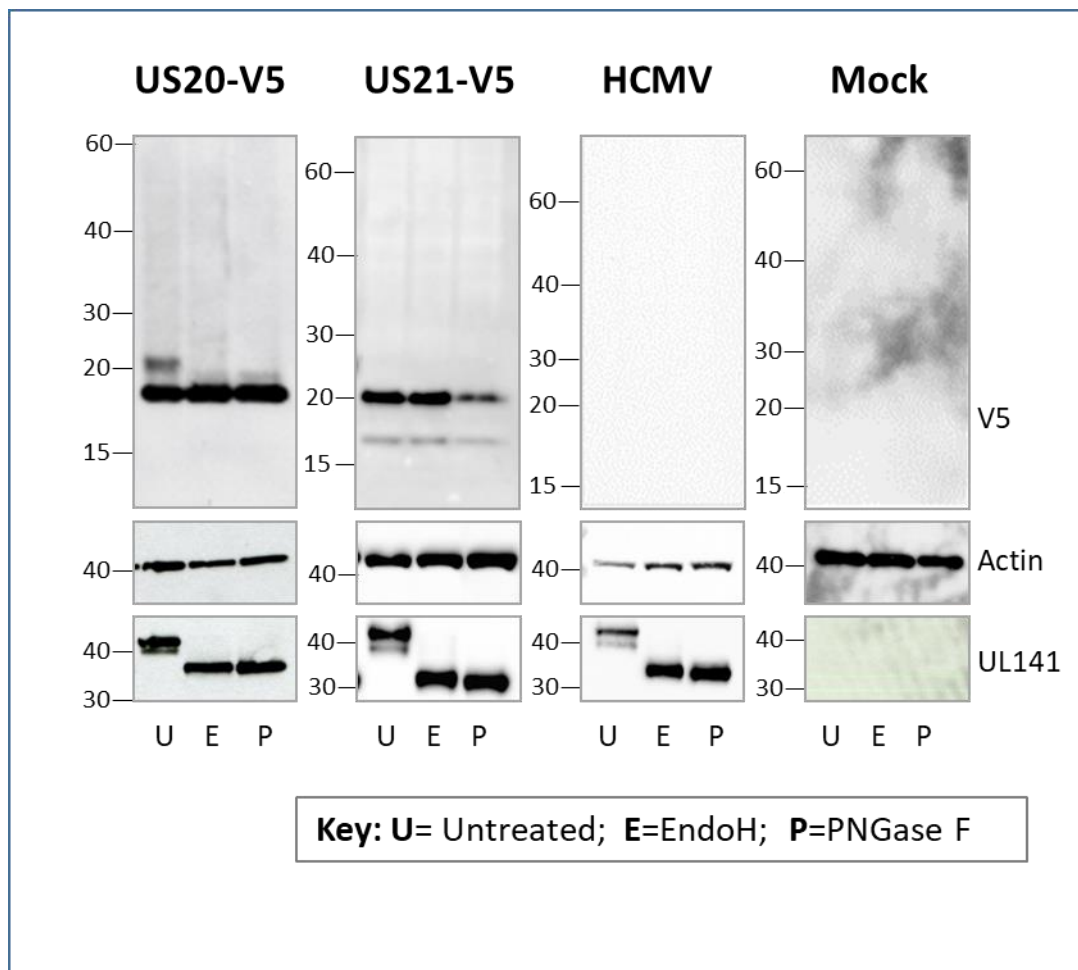
Figure 4.5: The N-glycosylation states of the US12 family proteins. Fibroblasts (HF-TERTs) were mock infected or infected with HCMV Merlin or HCMVs encoding V5-tagged US12 gene viruses at an MOI of 10 and collected at 72 hpi. Samples were processed using the Whole Cell Lysate (WCL) method and underwent overnight de-glycosylation with enzymes EndoH and PNGase F. Samples were detected using an anti-V5 antibody and an HRP secondary antibody, with anti-actin as a loading control and anti-UL141 as a positive control that is known to be N-glycosylated and EndoH sensitive. Removal of N-glycosylation can be seen by the reduction in the apparent molecular weight of the protein.

contain a useable N-glycosylation site that did not show signs of N-glycosylation by immunoblot. If the C-termini of US16 is non-cytoplasmic as predicted (**Table 4.3**), then according to the topology, the N-glycosylation site therefore would not be luminal (**Figure 4.1**) and therefore could not be glycosylated. For the same reason, only 2 of the 3 asparagine residues of US20 would be luminal and available for N-glycosylation in this orientation. Although no other reductions in molecular weight could be seen in other members, there were some slight differences seen between the treatment types, as discussed further in **Section 4.5**. With pUS12-V5, a considerable reduction in protein amount could be detected in the PNGase F treated sample, despite relatively consistent loading and the protein appearing to remain the same molecular weight as the untreated and EndoH treated samples (**Fig. 4.5**). Similar minor differences were also seen in the PNGase F treated sample of pUS13-V5 and pUS21-V5. With pUS13-V5 it was hard to distinguish whether there were any changes in band size, as although there seems to be a slight difference with the PNGase F treatment sample, there was no reduction in band size detectable to suggest N-glycosylation. Slight differences were also detected in the PNGase treated samples of pUS17-V5 and pUS18-V5 (fainter) and pUS15-V5 (higher detection) but these were likely to be due solely to loading differences. The apparent molecular weights of the US12 family are presented later in **Table 4.5 (Section 4.5)**.

4.5 A comparison of US12 family members when expressed in isolation and in HCMV expression systems

All US12 family members could be detected by immunoblotting when expressed from HCMV, however not all could be detected when expressed in isolation using the adenovirus vector expression system (**Section 4.3 and 4.4**). This was especially obvious in the untreated and EndoH treated samples adenovirus samples of pUS15-V5 and pUS18-V5, with Ad pUS20-V5 also demonstrating much weaker detection unless treated with PNGase F (**Fig. 4.4**). The differences seen were likely due to the harshness of the PNGase F treatment conditions which may have affected denaturation, or may potentially have caused easier access of the antibody to the tag. This also implies that the severe treatment conditions associated with PNGase F may be required just to get some of them into solution (as they are not detected without this treatment), and this may be due to their transmembrane structure. As this happened to a lesser extent in the HCMV samples, this

further solidified the importance of examining these proteins within the context of infection.

In both expression systems, US12 family members appear to have 1 or 2 predominant protein forms, the molecular weights of which were comparable across expression systems, with the greatest size difference seen for the main protein form of pUS12-V5, presenting as 43 kDa in isolation and 25 kDa in the context of HCMV infection (**Table 4.5**). Extra bands and/or extra 'smears' were seen for the majority of US12 family members, especially at higher exposure levels (**Fig. 3.8, Chapter 3**), but expressing them in isolation resulted in further protein forms and more prominent smearing patterns by immunoblot (**Fig. 4.4, 4.4, Table 4.5**). These multiple extra protein species were particularly prominent for Ad pUS14-V5, pUS15-V5 pUS16, pUS17-V5, pUS18 and pUS19 compared to their HCMV counterparts. Differences were also noted in the N-glycosylation patterns between the 2 systems, with the adenovirus expression system persistently demonstrating less N-glycosylation. Primarily, pUS20-V5 was shown to be N-glycosylated exclusively in the context of HCMV infection (**Fig. 4.5**). Although both systems demonstrated that pUS14 and pUS17 were partially N-glycosylated, differences were seen in their glycoforms, with more substantial protein subsets N-glycosylated in HCMV infection (**Fig. 4.4 and 4.4**). Differences in N-glycosylation of HCMV proteins between HCMV and Ad is a common occurrence (Griffin et al., 2005).

The US12 family molecular weights and banding patterns were much more consistent across HCMV experiments, however they did vary slightly between immunoblots so consensus or average sizes were calculated from across all immunoblot experiments (**Table 4.6**). The molecular weights of their main protein forms ranged from 17 to 38 kDa, and these were all within 10 kDa of their predicted sizes (**Table 4.5**). The slight variations seen were likely due to the slight differences in the way that the gels ran during immunoblotting, as well as the distance between the protein of interest and the size markers which may affect measurement accuracy.

4.6 Conclusions

The differences between the US12 family proteins expressed in isolation and expressed in the context of HCMV infection, could be caused by differential post-translational modifications such as ubiquitination, or from aberrant migration due to the proteins not being fully denatured. One possibility for the additional multiple higher bands observed in

Table 4.5: Summary of US12 family member protein sizes in comparison to their predicted sizes, their sizes in an adenovirus expression system, and their previous published sizes

Gene	Predicted size †	Average protein size in HCMV infection*	Protein size in adenovirus expression system	Protein sizes in published data
US12	32.48 kDa	25 kDa	~43 kDa	None
US13	29.46 kDa	21 kDa with smear >80 kDa	Ladder/smear	None
US14	34.3 kDa	37 kDa	Doublet of ~22 and 32kDa	None
US15	29.1 kDa	~23 kDa	~24 and ~40 kDa under PNGase treatment◊	None
US16	34.69 kDa	~28 kDa	~30 kDa	33 kDa (Bronzini et al., 2012)
US17	31.91 kDa	~26 kDa and faint band 23 kDa	~20 kDa, with fainter bands seen at ~24, 39 and 52 kDa	10 and 80 kDa (Das et al., 2006)
US18	30.2 kDa	~22 and 24 kDa	Doublet of ~18 and ~19 kDa under PNGase treatment◊	36 kDa (Guo)
US19	26.42 kDa	~19 kDa	~19 kDa, with possible higher bands	32 kDa (Guo)
US20	28.53 kDa	~19 and 23 kDa	~19, ~22 and 33 kDa	43 and 36 kDa (Guo); 25 and 30kDa (Cavaletto et al., 2015)
US21	26.93 kDa	~17 and 19 kDa	~17 and 17.5 kDa	None

† Predicted using Protein Molecular Weight prediction software at http://www.bioinformatics.org/sms/prot_mw.html

* A consensus taken from all immunoblots of the US12 family **(Table 4.6),** and rounded to the nearest kDa

◊ Could not be detected in untreated conditions

Table 4.6: Summary of HCMV US12 family member protein sizes across multiple immunoblots to determine an average or consensus molecular weight

Protein	Protein sizes◊				
	Figure 3.7	**Figure 4.2**	**Figure 4.5**	**Figure 5.1†**	**Figure 5.9 (in virion)**
US12	~24 kDa	~22 kDa	~27 kDa	28/29 kDa	25/26 kDa in virion
US13	21 kDa and smear to >80 kDa	Smear from ~22 to >160 kDa	Smear from 20 to >160 kDa	Smear from 21/22 to >60 kDa	~22 kDa in virion (and possible smear)
US14	~37 kda	~34 kDa	37/38 kDa (32/33 kDa de-glycosylated)	37/38 kDa	ND*
US15	20 kDa	~26/27 kDa	~22 kDa	~24 kDa	ND*
US16	28/29 kDa	~31 kDa	~27/28	25/26 kDa	27/28 kDa in virion
US17	~26 kDa and faint band a 22 kDa	~24/25 kDa and faint band at 22/23 kDa	28/29 kDa and ~25kDa (~25 kDa de-glycosylated)	~24 kDa and faint band at 21/22 kDa	ND*
US18	~22 and 25 kDa	~24/25 and 26/27 kDa	~20 and 22 kDa	~21 and 23/24 kDa	19, 20/21 and a faint band at 17 kDa in virion
US19	~19 kDa	~18 kDa	~20 kDa	19/20 kDa	~18 kDa in virion
US20	19 and ~24 kDa	18/19 and 22/23 kDa	18/19 and 21 kDa	~20 and 24/25 kDa	18/19 and ~22 kDa, and extra higher weight bands in the virion
US21	~16 and 19 kDa	~17 and 19 kDa	16/17 and 20 kDa	16 and 18/19 kDa	17/18 and 14 kDa in virion

◊ Protein sizes were calculated from the middle of the protein band. The extra bands detected at higher exposures have not been included as these were not detected or well-defined under "normal" conditions and were not consistently seen.
† Extra bands and extra smears/ladders were especially detected with the addition of leupeptin (a lysosomal inhibitor), especially for US12, US15 and US20
* ND= Not detected in the virion.

the adenovirus samples could be due to dimerisation or multimerisation of the protein which could possibly be present due to inefficient denaturing or due to the strong hydrophobic interactions within these proteins. This could be especially possible for Ad pUS12-V5 and pUS17-V5 which were detected at regular intervals that could represent monomers, dimers and multimers. Ad pUS12-V5 for example, could be seen around 20,40, 60 and 80 kDa (22, 43, 59 and 78 kDa) under PNGase treatment and Ad pUS17-V5 could be seen at around both 20 and 40 (~39) kDa (**Fig. 4.4**). The dissimilarities observed suggested that studying the US12 family within HCMV was much more beneficial and representative of infection and it appears that the family are not processed properly when expressed in isolation. It is therefore likely that they need other HCMV proteins to be processed correctly and possibly also to function, and they may require other US12 family members if they form complexes as hypothesised. This could explain why previous studies found little function when the proteins were expressed individually. It would be interesting to see whether expressing the whole family from an adenovirus vector would alleviate some of these problems.

Of the three US12 family members N-glycosylated in the context of HCMV infection, pUS20 was the only member previously demonstrated to be partially glycosylated in the HCMV strain TR (Cavaletto et al., 2015), with N-glycosylation a novel characteristic for pUS14 and pUS17. The removal of N-glycosylation from these 3 proteins by both EndoH and PNGase F suggests that all 3 of these proteins had not been processed through the Golgi. The significance of this N-glycosylation is discussed further in **Section 6.1**. For pUS17 and pUS20 which were only partially N-glycosylated, a proportion of these protein subsets therefore remained un-N-glycosylated during infection. One possibility is that their 2 protein species are in different sub-compartments, with UL141 previously demonstrated to be differentially glycosylated in the virion than it is within the cell (Dan Cochrane, thesis).

N-glycosylation was the sole cause of the pUS14-V5 increased molecular weight observed by immunoblot compared to its predicted size (**Table 4.5**) and pUS14-V5 was the only US12 family member whose main protein form appeared at a greater molecular weight than was predicted. All other members are likely to have alternative PTMs to account for their differences in size. Many of the other US12 family members were instead smaller than predicted (**Table 4.5**) which is generally more unusual. This can be more common for membrane proteins however as their hydrophobic residues within transmembrane domains can affect detergent binding and folding, which can lead to differential gel migration for the same size protein (Rath et al., 2009). The process of running

transmembrane proteins by SDS-PAGE could also be the cause of the 'smears' seen by immunoblot. Smaller-than-predicted molecular weights could also be caused by alternate splicing or cleavage for example.

As N-glycosylation was not the cause of the doublet bands of pUS18-V5 and pUS21-V5, alternative post-translational modifications such as phosphorylation, prenylation or methylation may be the explanation, else they may have undergone processing such as alternative splicing or cleavage. For Merlin US18, a shorter transcript isoform had also been identified in which transcription was initiated within the US18 coding sequence and translation was predicted to produce an N-terminally truncated version of pUS18 (Stern-Ginossar et al., 2012). However, this RNA transcript was first detected until 72 hpi, whereas our doublet band could be observed from 48 hpi. This truncated 350 bp RNA transcript was also predicted to encode a 12.95 kDa product, smaller than the faster migrating species detected in this thesis (~20 kDa). Due to the differences in timing and size, it seems less likely that this transcript is the cause of the second protein form in pUS18-V5 and more likely to be caused by post-translational modifications.

US12 family member protein sizes were also relatively similar (although again, often smaller) to previously published data on US12 family members (**Table 4.5**). Differences seen could be due to the use of different HCMV strains and different cell types along with the different methods of detection used. The V5 tag is only predicted to add ~1.4 kDa onto the protein's molecular weight, but other tags used will alter a protein's molecular weight by differing amounts, so the use of alternative epitope tags may also have played a role in the differences seen. Despite most US12 family members appearing to be of a similar size across different strains, pUS17 was shown to have the biggest difference in size, with AD169 pUS17 detected at 10 and 80 kDa (**Table 4.5**). On the other hand, Merlin pUS17-V5 in this thesis was detected at ~23 and 26 kDa, which was much closer to the predicted 31.9 kDa size. Strain differences or post-translational modifications could be the cause of this variance and this would need to be investigated further. AD169 pUS17 was previously demonstrated to be segmented through the separate detection of its C- and N-termini (Das and Pellett, 2007, Das et al., 2006), and an N-terminal tag or antibody would be required to establish whether Merlin pUS17 was also segmented or not. N-terminal tags and/or specific US12 family member antibodies may be beneficial to help study the US12 family in the future, and in confirming the apparent detected size of each US12 family member protein.

Each member has been predicted to have 7TM domains, and in order for the N-glycosylation motif sites to be accessed, the asparagine residues of pUS14, pUS17 and pUS20 would need to be in the lumen of the ER. This would require each of their C-termini to be lumenal (non-cytoplasmic), in order for the majority of N-glycosylation sites to likewise being lumenal, and correspondingly, each US12 family member is generally predicted to have non-cytoplasmic C-terminals and cytoplasmic N-terminals (**Fig. 4.1, Table 4.3**). The implications of this topology are discussed further in **Section 6.1 and 6.3.1**.

Knowing their expression patterns may also aid in determining likely functions of the US12 family members. Immunoblotting appeared more sensitive and superior for detecting this gene family, detecting all 10 family members compared to only 6 by proteomics (Weekes et al., 2014). It is not known why the immunoblotting would be more sensitive than proteomic mass spectrometry in the case of the US12 family protein detection, but it may be due to the hydrophobic nature of the proteins due to the number of transmembrane domains, differences in sample processing or the low quantity of US12 family proteins expressed in the cell. Of the six US12 family members detected by proteomics, each was only detected by 1-3 peptides, and even the more sensitive detection system using the new Orbitrap Fusion in the proteomics WCL3 experiment only detected a maximum of 2 peptides per US12 family protein. This together with the fact that US12 was also the only US12 family member detected in all PM and WCL proteomics samples (Weekes et al., 2014), supports that low abundance could definitely be a contributing factor to the lack of detection by proteomics (Weekes et al., 2014). Other factors may include the use of mass spectrometry and software databases used to identify the proteins, as deducing chemical structures and subsequently protein sequences from mass-to-charge ratios can be a challenging computational task (Bruce et al., 2013). There is also the possibility that the V5 tag could be affecting the stability or turnover rates of the US12 family proteins, and this could explain why V5-tagged US12 family proteins were better detected in the immunoblots rather than in the proteomics in which un-tagged US12 family proteins were detected by mass spectrometry. A differing rate of protein degradation was previously demonstrated between tagged and untagged versions of human protein DJ-1, where the N-terminal Flag-tagged DJ-1 had a reduced rate of degradation compared to the C-terminal V5 tagged version or the untagged version, resulting in higher cellular levels (Alvarez-Castelao et al., 2012). Therefore there is a small chance that the degradation rates of V5-tagged US12 family members may differ slightly to those *in vivo*, however the use of tagged proteins is unavoidable at present due to the lack of specific US12 family antibodies.

The immunoblot time-course revealed that US12 family members ranged from early Tp2 proteins (such as pUS20) to late Tp5 proteins (such as pUS17), with all members detected by 48 hpi, and the majority of US12 family protein expression peaking around 72 hpi. As well as being broadly in agreement with the QTV data, these immunoblot expression patterns were also demonstrated to be in accordance with other limited published data available on US12 family kinetics (Bronzini et al., 2012, Guo, 1993, Luganini et al., 2017) (**Table 4.4**). pUS20 for example, was designated by immunoblot and QTV to be a Tp2 protein and this was also consistent with previous findings in which pUS20 was detected from 18 hpi until 72 hpi (Guo, 1993) or detected equally from 24 hpi to 96 hpi (Cavaletto et al., 2015).

These immunoblot data have advanced upon previous findings by having an improved detection of US12 family members that were previously either not detected, only partially detected, or those only detected in unquantified immunofluorescence studies. They have also enhanced preceding studies through the use of a more clinically relevant strain. The novel insight that the majority of the US12 family members are expressed as Tp3 proteins may help to determine at what point during infection they are exerting their function and may hint at their potential functions (**Discussion 6.1**).

Taking everything into account, the US12 family proteins were demonstrated to have incorrect processing outside of the context of HCMV infection, and all further experiments were consequently undertaken solely in HCMV. In conclusion, the US12 family findings corroborate with the limited data currently known, and novel characteristics have been revealed, increasing the knowledge on the basic fundamentals of the US12 family in regards to their expression kinetics, protein sizes and N-glycosylation states.

5: Localisation of the US12 family members within infected cells

An appreciation of a HCMV protein's localisation within a cell can inform on its role or function, with IE1's transient association with punctate nuclear domains previously helping to inform on its role in disrupting nuclear bodies (Wilkinson et al., 1998, Kelly et al., 1995). Nuclear localisation can often indicate regulatory or structural roles, with the regulatory protein IE2 and the structural major capsid protein (pUL86) also found within the nucleus (Sanchez et al., 1998). Tightly associated tegument proteins are most likely added to the capsid in the nucleus, whereas outer tegument proteins are often added later in the virion assembly compartment (vAC) (Sanchez et al., 2000b). The tegument protein pp65 (UL83) accumulates along the nuclear periphery, and was found to have a direct interaction with nuclear lamins, linking its location to its function (Sanchez et al., 1998). Therefore, tracking the location of specific US12 family members within HCMV infected cells was undertaken for the primary goal of characterising their expression, but with the further aspiration that it might provide further insights into their roles, mechanisms-of-action and interactions within the cell. Functional studies had revealed that the US12 family regulated the cell surface expression of a broad range of plasma membrane proteins involved in different aspects of the immune response, including the regulation of cell adhesion molecules, NK cell ligands, cytokines and cytokine receptors (Fielding et al., 2017). Moreover, it was established that many of these immune ligands were targeted by multiple members of the US12 family. This was most notable for US18 and US20 which acted in co-operation to target the NKG2D ligand MICA and the NKp30 ligand B7-H6, and both further targeted an additional set of proteins (Fielding et al., 2014, Fielding et al., 2017). This lead to the hypothesis that multiple US12 family members were forming and acting as complexes. In that context, a major objective was to determine whether multiple US12 family members trafficked to the same locations.

The localisations of US12 family members were investigated within the context of HCMV infection, as US12 family members had demonstrated not to be processed correctly in isolation (**Chapter 4**), and due to the virus re-organising the host organelles during infection (**Section 1.4.1.6**).

5.1 Rescue of US12 family member expression by the inhibitor leupeptin

Proteomic studies showed that many of the cellular targets of the US12 family could be rescued by the addition of leupeptin (Fielding et al., 2017), an inhibitor of serine and thiol proteases that suppresses lysosomal degradation by 80-85% (Grinde and Seglen, 1980).

This observation led us to believe that the US12 family may use lysosomal degradation as a mechanism to degrade their target proteins. I sought to investigate whether the US12 family were similarly degraded themselves and whether they could likewise be rescued by the inhibitor leupeptin. To this end, fibroblasts were infected with the HCMV library of V5-tagged US12 family members, with the leupeptin added before harvesting the cells. MICA was used as a positive control protein as it had previously been shown to be targeted for lysosomal degradation and rescued by leupeptin addition (Fielding et al., 2014, Fielding et al., 2017).

The changes in expression of US12 family member proteins upon the addition of leupeptin could be classified into 3 distinct groups. Firstly, a subset of US12 family members (pUS16-V5, pUS19-V5 and pUS21-V5) appeared to be unaffected, with their expression remaining consistent with or without leupeptin addition and thus appeared not to be degraded in lysosomes (**Fig. 5.1**). Secondly, there were US12 family members that showed a substantial increase in abundance, including an increase in their dominant protein form(s); as demonstrated by pUS12-V5, pUS13-V5, pUS15-V5, pUS18-V5 and gpUS20-V5 (**Fig. 5.1**). This increase was most pronounced with pUS12-V5 (~15-fold) or could be more subtle, such as with pUS18-V5 (~2-fold), as calculated by densitometry. Generally an increase in the abundance of the main form of the protein was observed alongside increases in higher molecular weight bands, with pUS18-V5 being the exception. For example, leupeptin treatment not only resulted in an increase in abundance of the main protein species of pUS15-V5 (~24 kDa), but also revealed a ladder/smear of multiple higher molecular weight species ranging up to nearly 60 kDa. The final subset of proteins did not appear to have an increase in expression levels of the main protein species, however an increase was observed in some of the higher molecular weight forms only. This was observed with gpUS14-V5 and gpUS17-V5 (**Fig. 5.1**), although this was more evident at higher exposure levels (**Fig. 5.2**). For gpUS14-V5, the main form of the protein was detected at 37/38 kDa and remained unchanged but an increase occurred in the bands/smear that ranged between 40 and 80 kDa (**Fig. 5.1 & 5.2**). Similarly, the main doublet of gpUS17-V5 remained unaffected by leupeptin, but an increase occurred in the smear at 28-35 kDa (**Fig. 5.1 & 5.2**). For both of these US12 family members, multiple additional protein forms were also revealed at these longer exposure times, however these also remained at equal levels of expression with and without leupeptin treatment (**Fig. 5.2**).

The three US12 family members whose expression were not modified by leupeptin (pUS16-V5, pUS19-V5 and pUS21-V5), were correspondingly not predicted to contain the

US12-V5   US13-V5   US14-V5   US15-V5

V5

Actin

MICA/
B

U   L   U   L   U   L   U   L

US16-V5   US17-V5   US18-V5   US19-V5

V5

Actin

MICA/
B

U   L   U   L   U   L   U   L

U= Untreated **L**= Leupeptin treatment

Figure 5.1

169
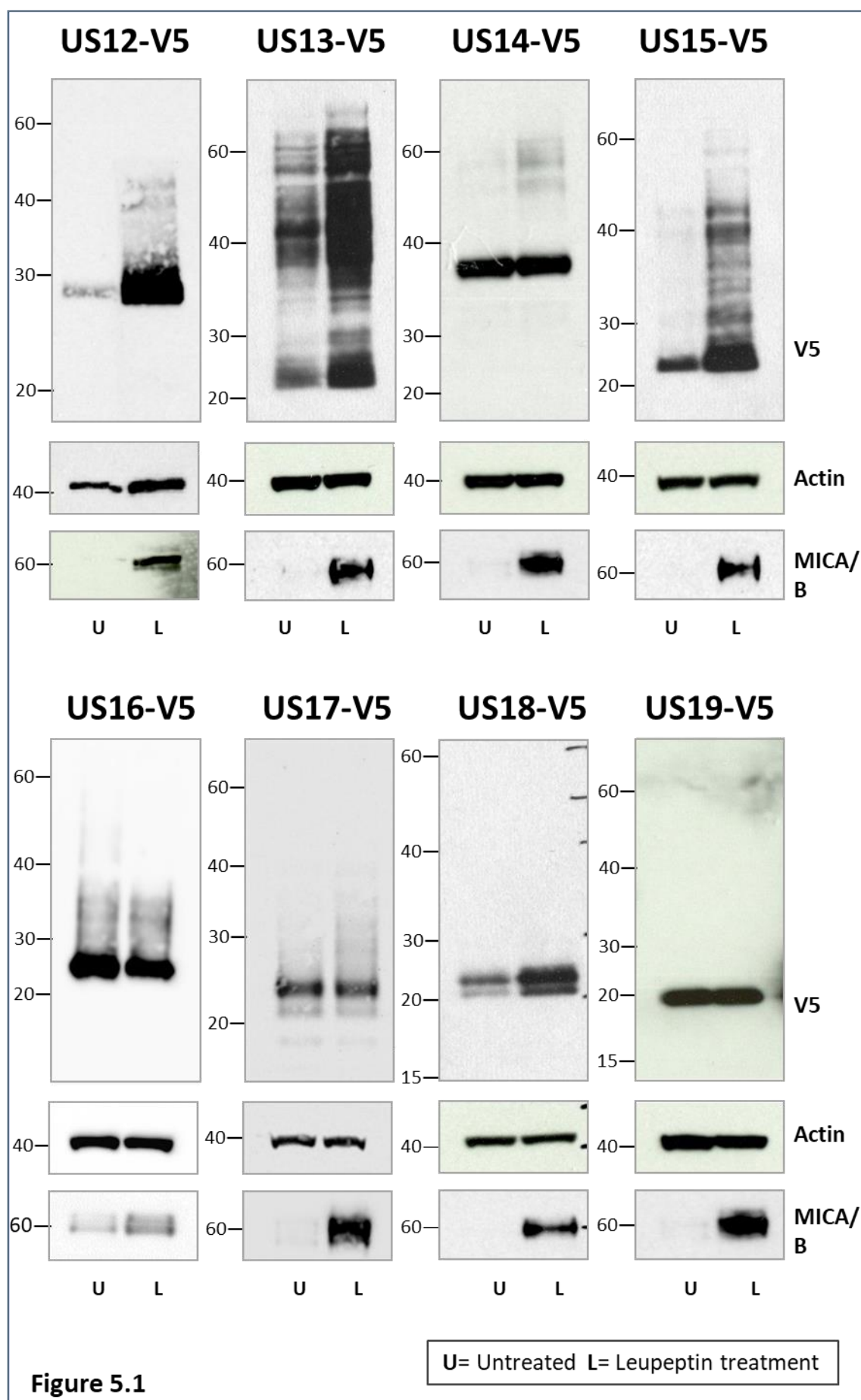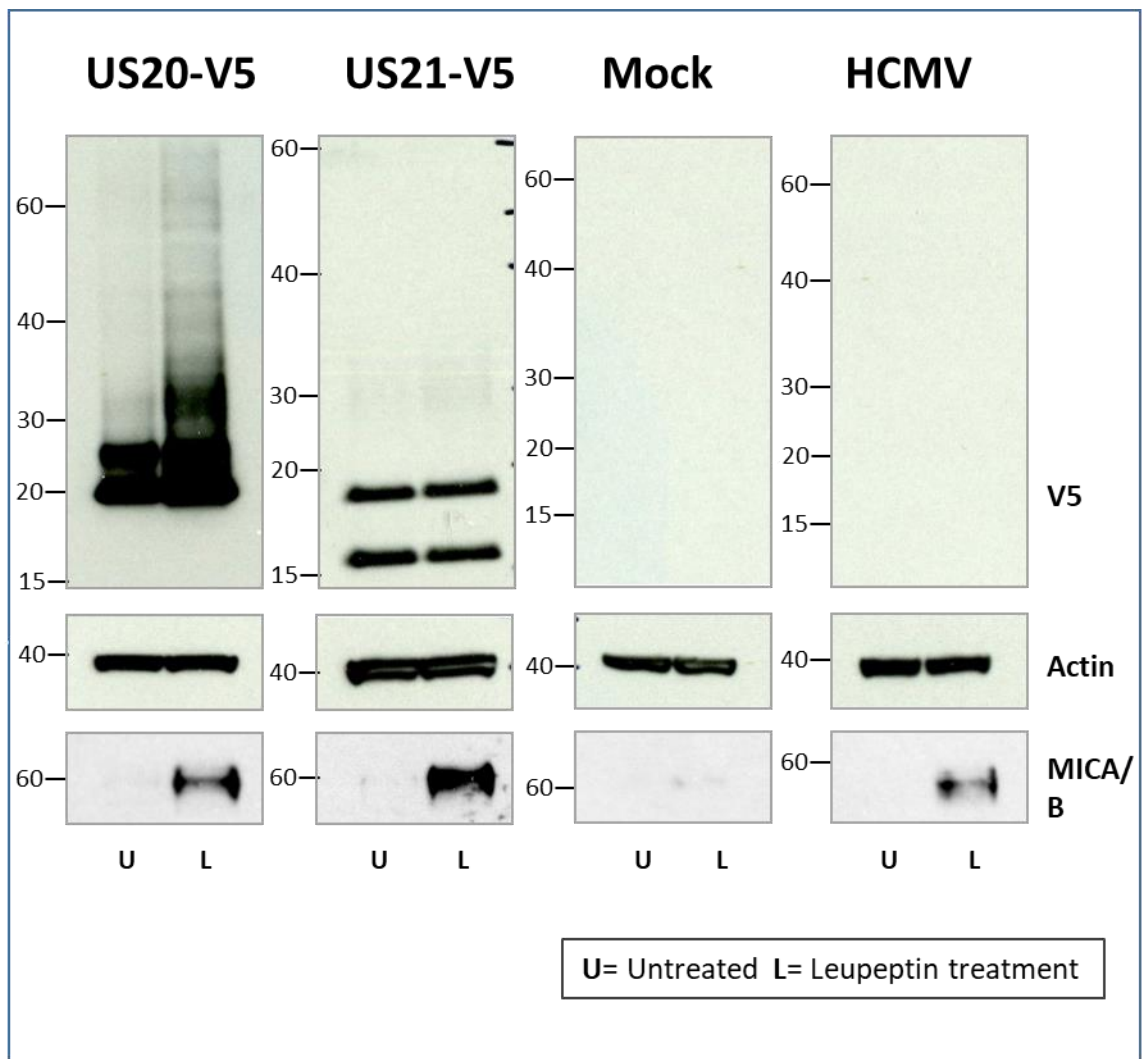
Figure 5.1: US12 family members are degraded and are rescued by the addition of the lysosomal inhibitor leupeptin. Fibroblasts were mock infected or infected with HCMV Merlin, or HCMVs encoding V5 tagged US12 genes at an MOI of 10. Samples were left untreated or had leupeptin (a lysosomal inhibitor) added 18 hours prior to harvesting at 72 hpi. Proteins were extracted and US12 family members detected using an anti-V5 antibody and an HRP secondary antibody. All samples were run with anti-actin used as a loading control & MICA as a positive control.
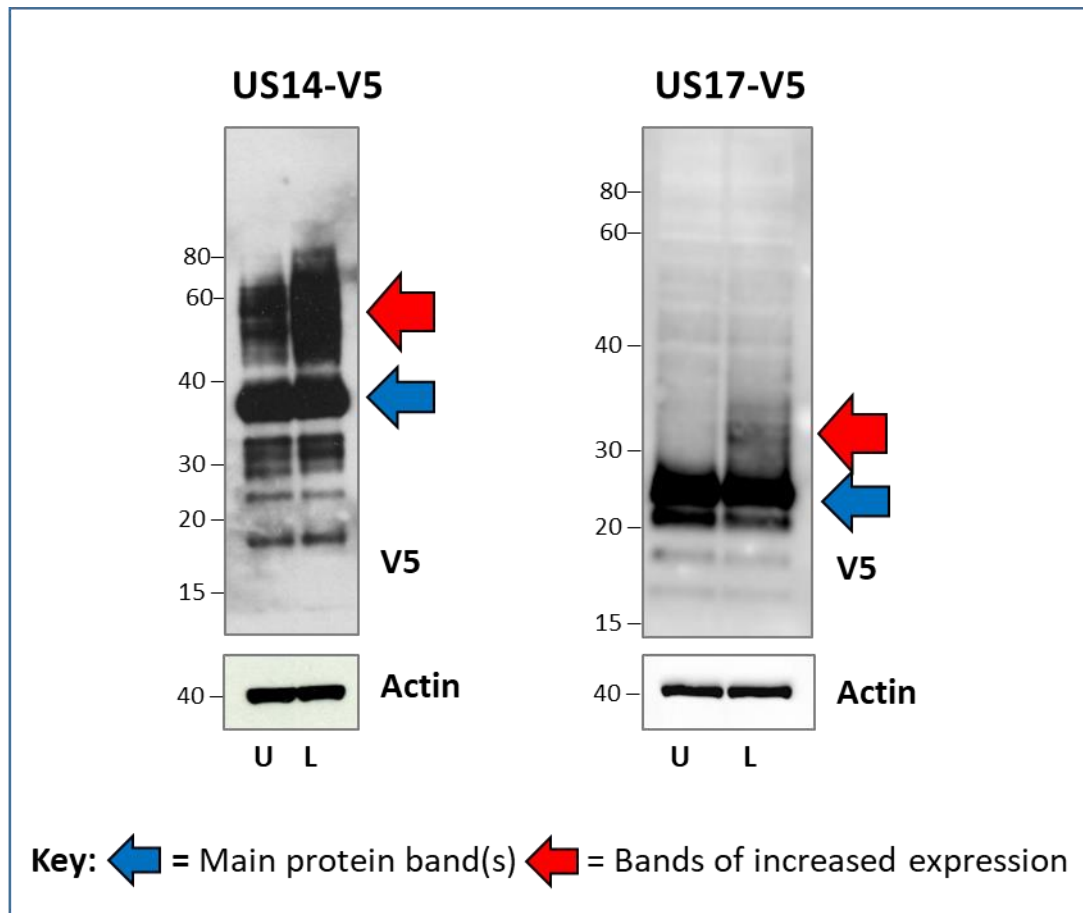
Figure 5.2: Increase in protein expression of higher weight molecular bands at higher exposure levels of pUS14-V5 and pUS17-V5. Fibroblasts were mock infected or infected with HCMV Merlin, or HCMVs encoding V5 tagged US12 genes at an MOI of 10. Samples were left untreated or had leupeptin (a lysosomal inhibitor) added 18 hours prior to harvesting at 72h.p.i. Proteins were extracted and detected using an anti-V5 antibody and an HRP secondary antibody. All samples were run with anti-actin used as a loading control. Areas of protein bands of increased expression highlighted by red arrows, and the 'main' bands seen at normal exposure levels indicate by blue arrows.

TRG_LysEnd_APsAcLL_1 lysosomal targeting sequences by Eukaryotic Linear motif (ELM) software as previously described (**Section 4.1.2, Table 4.2)**. Of the 5 family members shown to have their dominant protein forms regulated by leupeptin, US12, US13, US18 and US20 were all predicted to contain 1 or 2 lysosomal targeting sites from their sequence which showed a good correlation between the predicted and experimental results (**Section 4.1.2).** The only member for which an increase in protein abundance was detected without the identification of the TRG_LysEnd_APsAcLL_1 motif was pUS15, along with the higher molecular weight forms of gpUS17 (**Section 4.1.2)**.

These results demonstrated that leupeptin inhibited the degradation of the main protein form(s) of 5 family members and rescued higher molecular weight forms of 6 members. This result is consistent with the observation that the US12 family is involved in targeting proteins for proteolysis and indicates that these viral proteins may be co-degraded alongside their targets, potentially in lysosomes.

5.2 Immunofluorescence co-localisation of US12 family members with lysosomes

Since the expression of six US12 family members could be rescued by leupeptin, it was a natural next step to investigate whether they co-localised with lysosomes. Fibroblasts infected with V5-tagged US12 family members were incubated with or without leupeptin for 18 h prior to cells being stained with LysoTracker, a red fluorescent dye which labels lysosomes. Cells were then fixed, permeabilised and stained with anti-V5 antibodies to detect the individual US12 family members.

The lysosomes were first investigated for their localisation within the cell in comparison to the virion assembly compartment (vAC), which is found adjacent to the nucleus and defined by pp28 staining (**Section 1.4.1.6**). In HCMV infected cells the lysosomes were organised in a ring/zone around the vAC with their association demonstrated in **Figure 5.3.** In uninfected cells, the lysosomes instead appeared to be distributed throughout the cytoplasm of the cells (**Fig. 5.4**).

No member of the US12 family co-localises specifically with the lysosomes even in the presence of leupeptin, as assessed by overlap with the LysoTracker dye (**Fig. 5.4**). Although the US12 family proteins did not generally exhibit evidence of specific co-localisation with the lysosomes, they were often found in the same region of the cell as the lysosomes,

Key:
pp28 (vAC, green),
LysoTracker
(lysosomes, red),
DAPI nuclear stain
(blue)
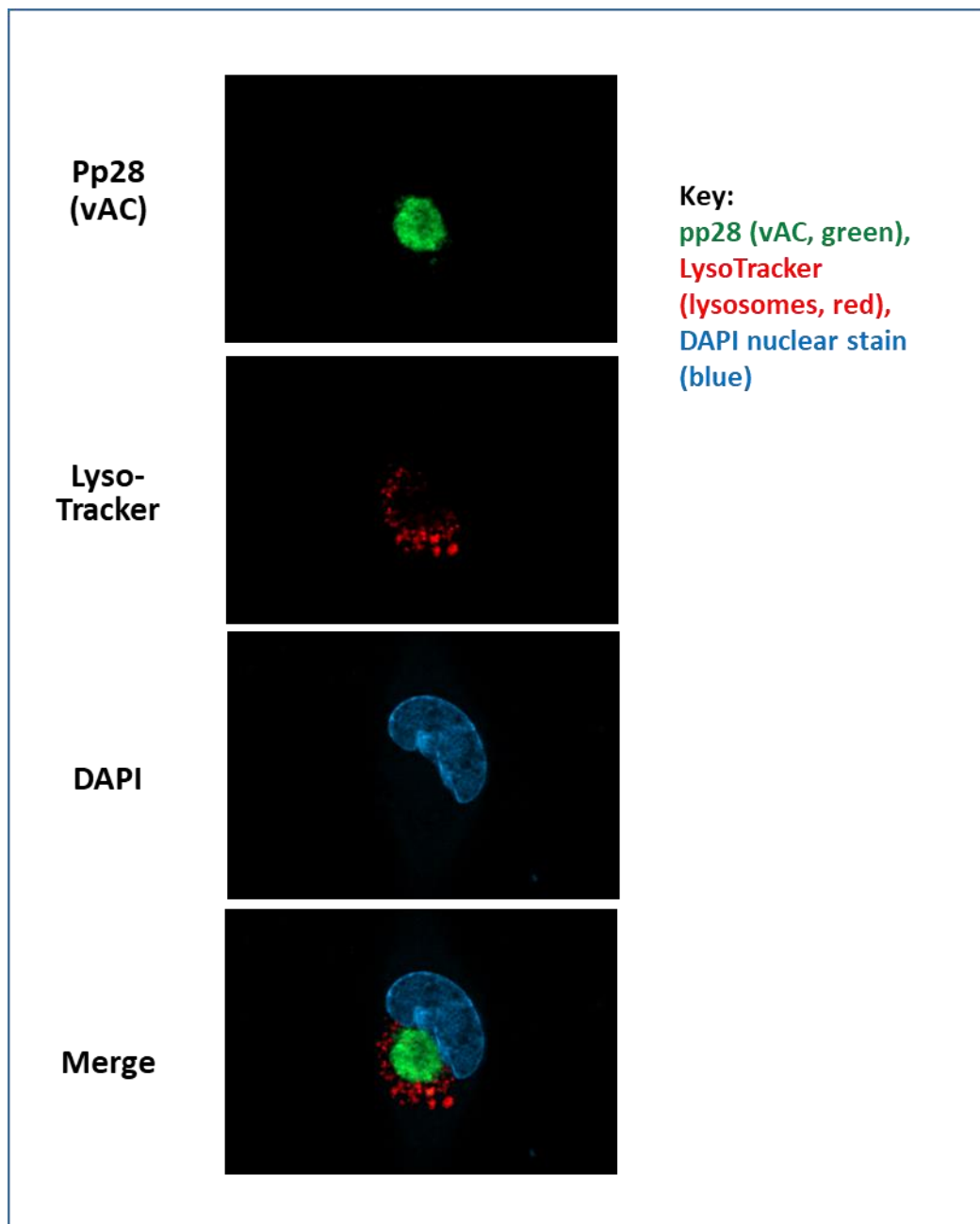
Figure 5.3: The association of the vAC marker pp28 with lysosomes. In relation to **Figure 5.4**, fibroblasts (HF-TERTs) were infected with HCMV Merlin at MOI of 10. Samples were left untreated and lysosomes were stained with LysoTracker (red) at 72 h.p.i. and then fixed. Cells were then co-stained with the anti-pp28 antibody (mouse) and anti-mouse Alexa-Fluor AF488 (green), with the nucleus identified with the nuclear stain DAPI. Images were taken using a Zeiss microscope (Axio Observer Z1 with ApoTome) using optional sectioning.

US12-V5

Figure 5.4

US13-V5

- Leupeptin **+ Leupeptin**

V5

Lyso-Tracker

DAPI

Merge

Key: V5 tag (green),
LysoTracker (lysosomes, red),
DAPI nuclear stain (blue)

Figure 5.4 cont.

US14-V5

	- Leupeptin	+ Leupeptin

V5

Lyso-Tracker

DAPI

Merge

Key: V5 tag (green), LysoTracker (lysosomes, red), DAPI nuclear stain (blue)

Figure 5.4 cont.

US15-V5

Key: V5 tag (green),
LysoTracker (lysosomes, red),
DAPI nuclear stain (blue)

Figure 5.4 cont.

US16-V5

	- Leupeptin	+ Leupeptin
V5		
Lyso-Tracker		
DAPI		
Merge		

Key: V5 tag (green),
LysoTracker (lysosomes, red),
DAPI nuclear stain (blue)

Figure 5.4 cont.

178

US17-V5

Figure 5.4 cont.

Key: V5 tag (green), LysoTracker (lysosomes, red), DAPI nuclear stain (blue)

US18-V5

Figure 5.4 cont.

Key: V5 tag (green),
LysoTracker (lysosomes, red),
DAPI nuclear stain (blue)

US19-V5

Key: V5 tag (green), LysoTracker (lysosomes, red), DAPI nuclear stain (blue)

Figure 5.4 cont.

US20-V5

Key: V5 tag (green), LysoTracker (lysosomes, red), DAPI nuclear stain (blue)

Figure 5.4 cont.

US21-V5

	- Leupeptin	+ Leupeptin
V5		
Lyso-Tracker		
DAPI		
Merge		

Key: V5 tag (green),
LysoTracker (lysosomes, red),
DAPI nuclear stain (blue)

Figure 5.4 cont.

HCMV 1111

	- Leupeptin	+ Leupeptin
V5		
Lyso-Tracker		
DAPI		
Merge		

Key: V5 tag (green), LysoTracker (lysosomes, red), DAPI nuclear stain (blue)

Figure 5.4 cont.

Key: V5 tag (green), LysoTracker (lysosomes, red), DAPI nuclear stain (blue)

Figure 5.4: The US12 family member proteins are not generally associated with lysosomes. Fibroblasts (HF-TERTs) were mock infected or infected with HCMV Merlin or HCMVs encoding V5-tagged US12 gene viruses at MOI of 10. Samples were either left untreated or had leupeptin (a lysosomal inhibitor) added 18 hours prior to harvesting. Lysosomes were stained with LysoTracker (red) at 72 h.p.i. and then fixed. Cells were then co-stained with the anti-V5 antibody (mouse) and anti-mouse Alexa-Fluor AF488 (green), with the nucleus identified with the nuclear stain DAPI. Images were taken using a Zeiss microscope (Axio Observer Z1 with ApoTome) using optional sectioning. Images representative of 2 independent experiments.

which occurred with pUS12-V5, pUS13-V5, gpUS14-V5, pUS16-V5, gpUS17-V5, pUS18-V5 and pUS21-V5 (**Fig. 5.4**). Limited evidence of selective co-localisation with lysosomes (yellow fluorescence staining) was discernible for some family members including pUS13-V5, pUS15-V5 and gpUS20-V5 upon the addition of leupeptin as demonstra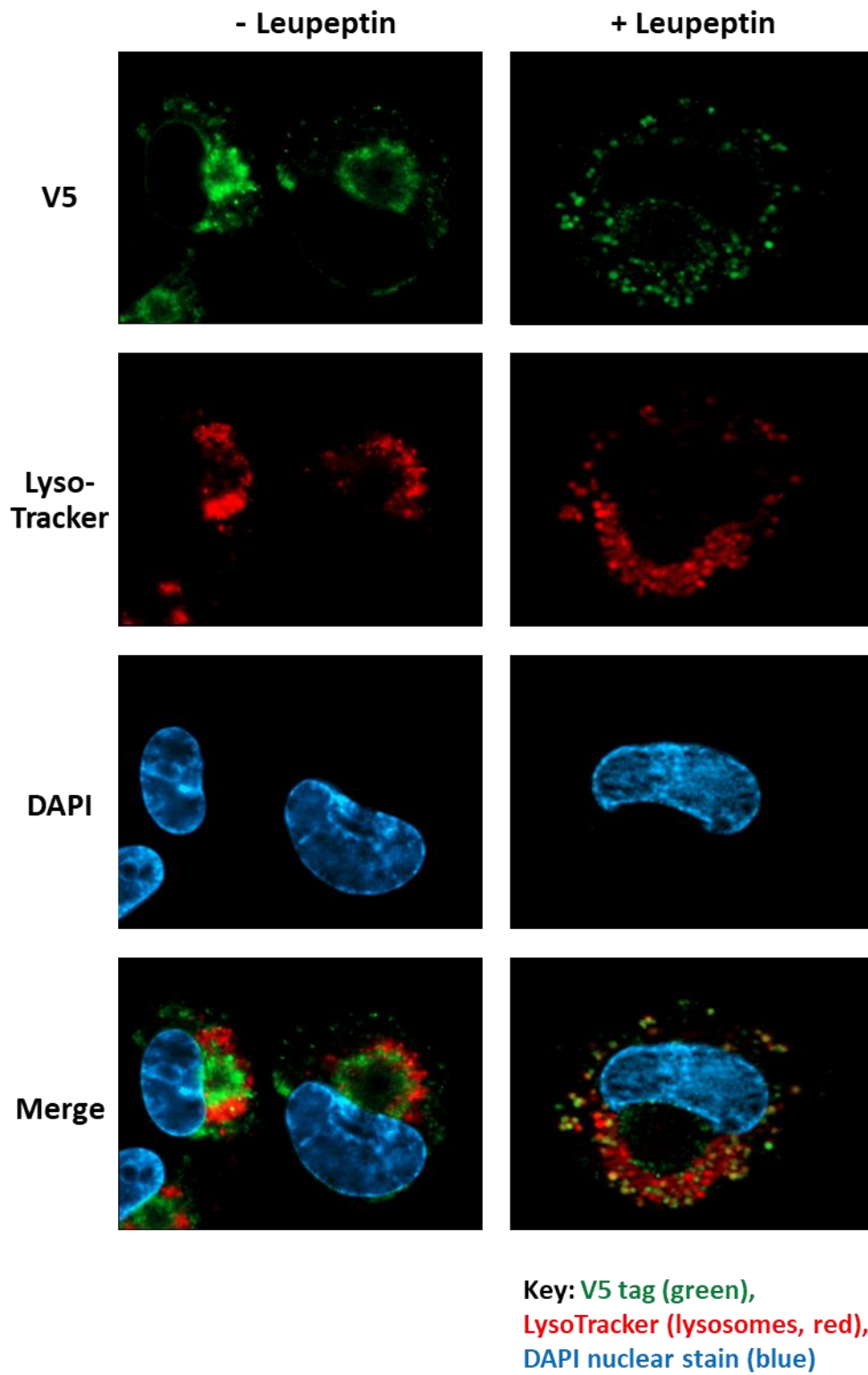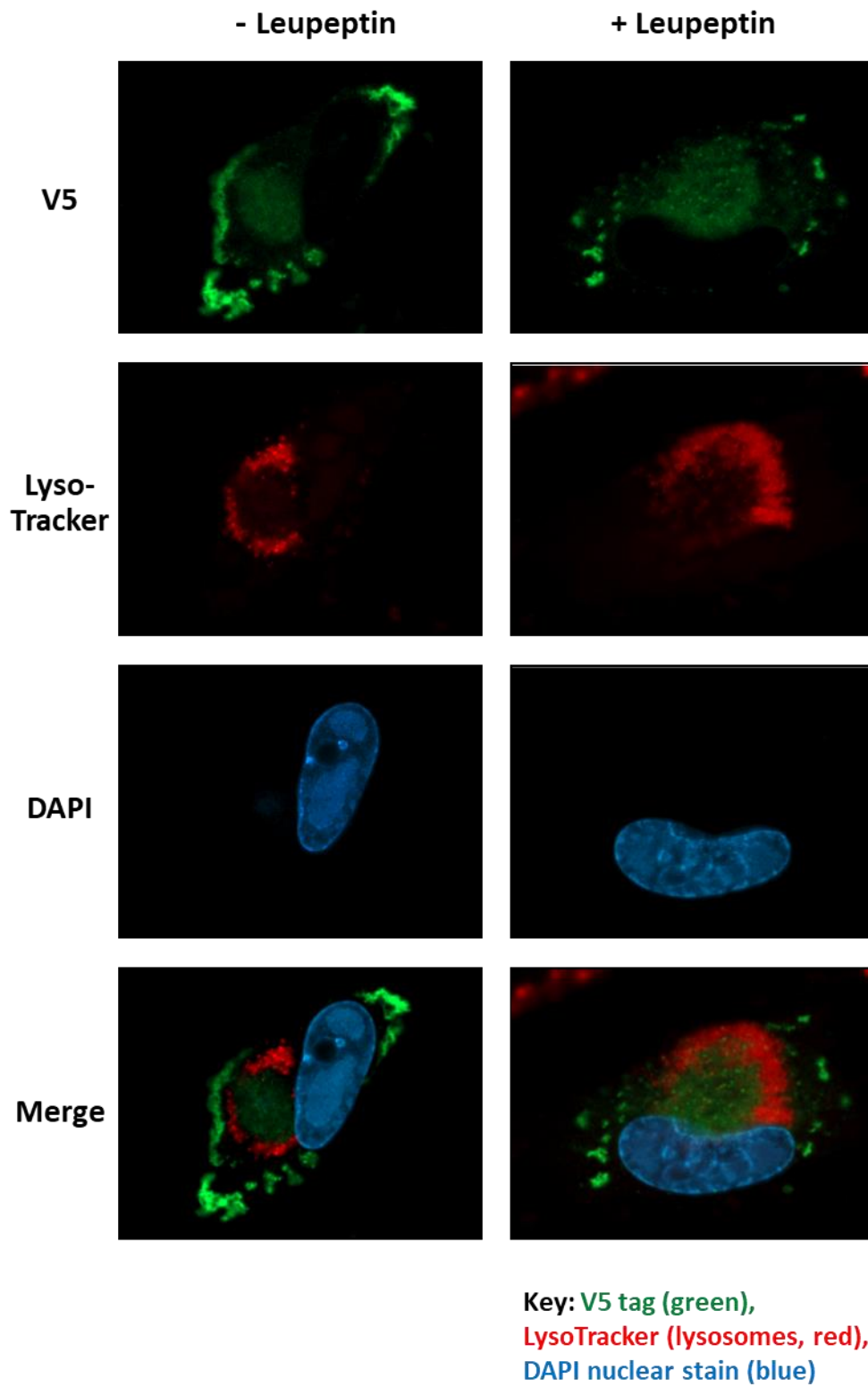ted in **Figure 5.5**. This co-localisation correlated to these 3 members being rescued by the addition of leupeptin by immunoblot (**Fig. 5.1**). Overall, these data suggested that the US12 family do not commonly co-localise with lysosomes *in vitro*. The addition of leupeptin does not appear to significantly change the US12 family's localisations within the cell, in which they are detected in juxta-nuclear cytoplasmic inclusions and/or sub-cellular punctate structure locations that required further investigation.

5.3 Immunofluorescence co-localisation of US12 family members with the endoplasmic reticulum

While US12 family members did not appear to typically traffic to lysosomes, a number were detected in defined sub-cellular punctate structures within the cell. To investigate whether these structures were the endoplasmic reticulum (ER), the US12 family members were co-stained with the ER marker calnexin at 72 hpi in fibroblasts. Although calnexin is an ER protein, it should be noted that HCMV re-models host cell organelles during productive infection to promote virus replication and maturation, best illustrated by the formation of the vAC which includes the rearrangement of organelles such as the Golgi, TGN and early endosomes (**Section 1.4.1.6**). The ER is thus not the same in infected cells as it is prior to infection as "ER-derived" membranous structures are being hijacked by the virus.

The sub-cellular punctate structures (SPS) described above was identified as the ER for pUS12-V5, pUS13-V5, gpUS14-V5, pUS15-V5, pUS16-V5, gpUS17-V5, gpUS20-V5 and pUS21-V5 (Fig 5.2) (**Fig. 5.6**). pUS18-V5 appeared mostly cytoplasmic with only a very faint concentration with the ER marker calnexin. Correspondingly, all of these US12 family members were identified as having the predicted ER retention motif 'TRG_ER_diArg_1', except for US18 and US21 as predicted by the Eukaryotic Linear Motif (ELM) software as described in **Section 4.1.2 (Table 4.2)**. US21 however was predicted to possess an XXRR-like ER retention motif by Psort II software (**Section 4.1.2**). Therefore pUS18-V5 was the only member that appeared to faintly co-localise with the ER despite not containing an ER motif, and pUS19-V5 was the only member predicted to contain an ER retention motif that was not localised to the ER.

Figure 5.5

US20-V5 +Leu

V5

Lyso-Tracker

DAPI

Merge

Key:
V5 tag (green),
LysoTracker
(lysosomes, red),
DAPI nuclear stain
(blue)
Yellow indicates
crossover of the V5-
tagged protein with
the lysosomes
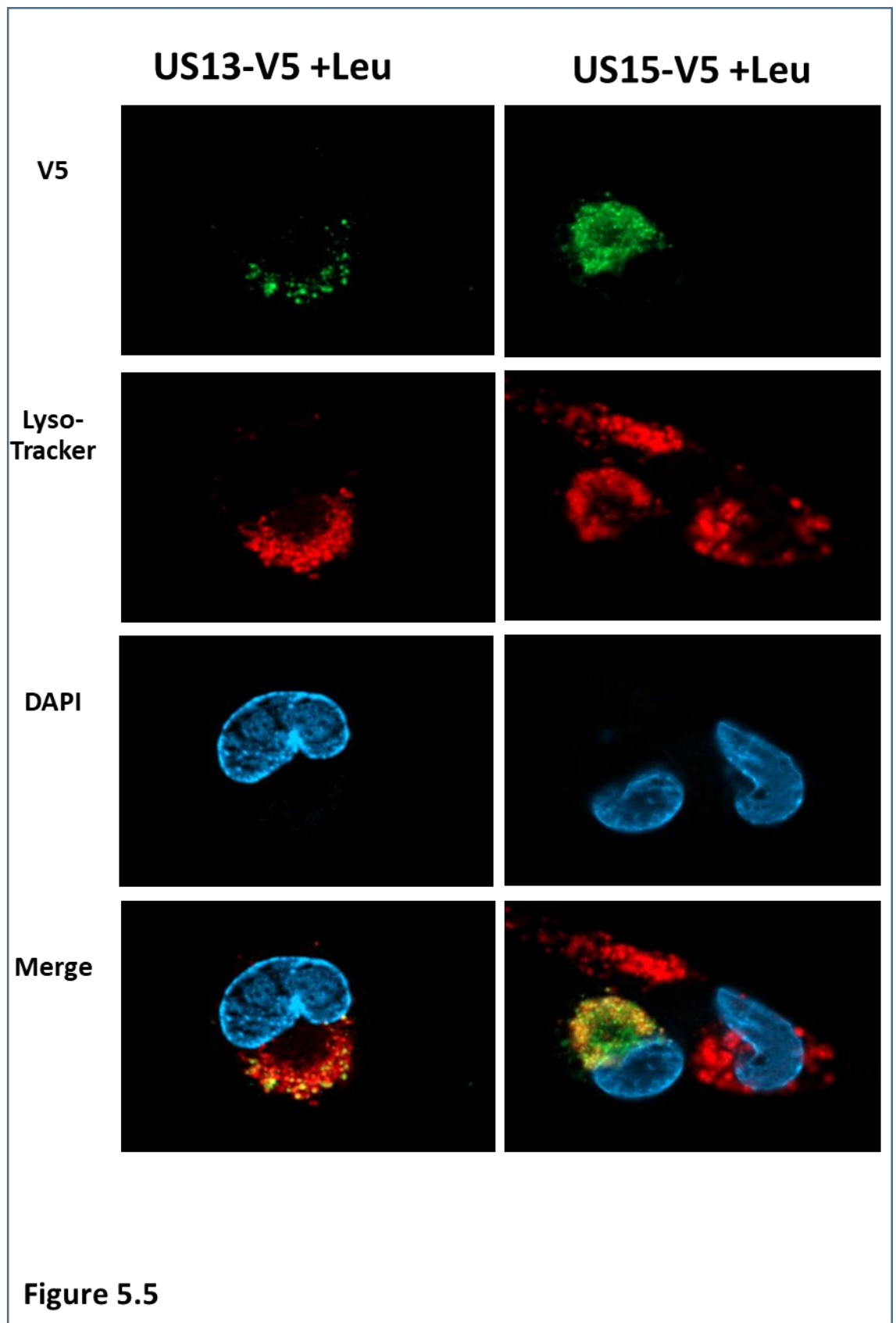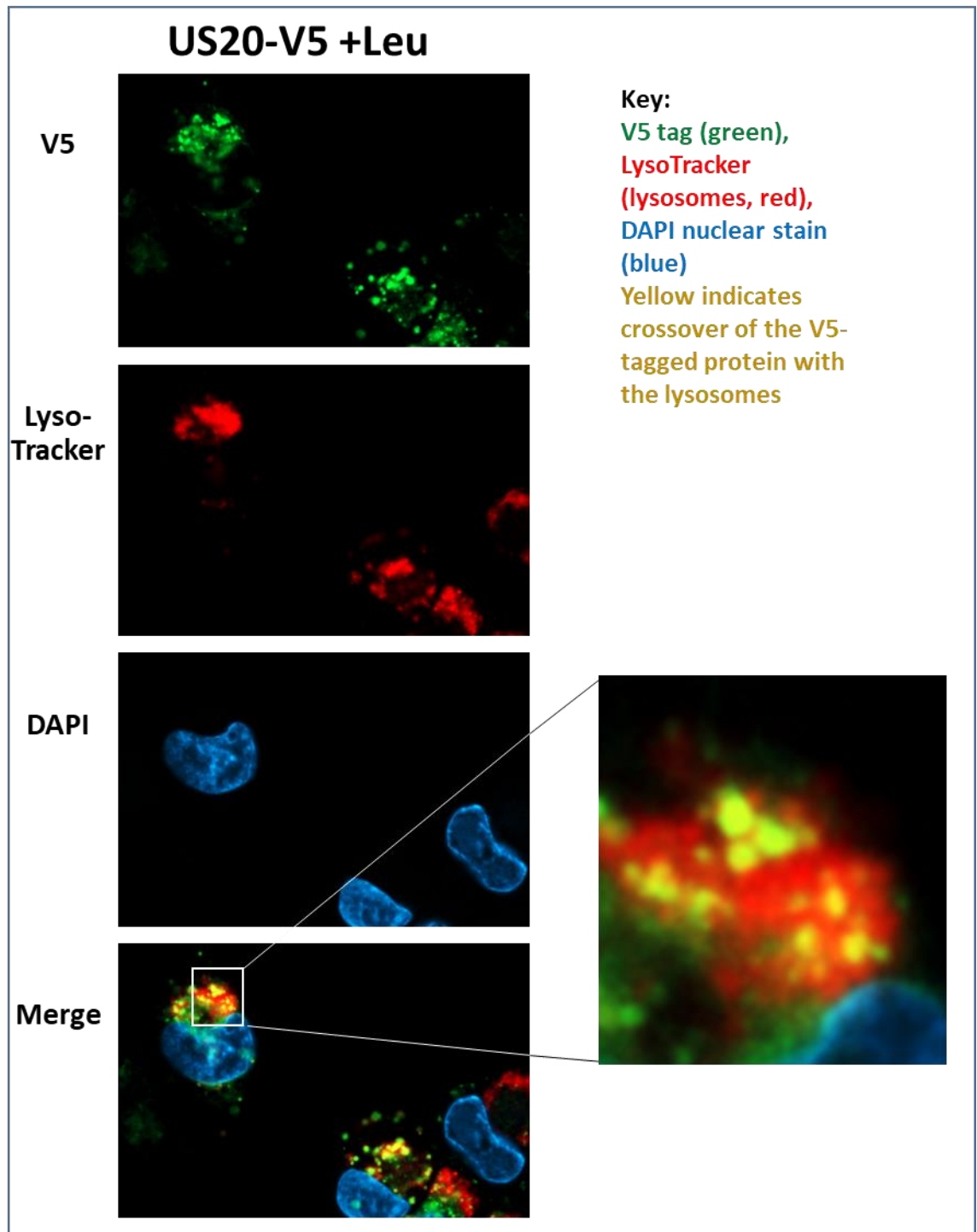
Figure 5.5 Rare examples in which US12 family member proteins show some association with lysosomes. Extra images in relation to **Figure 5.2**, showing some partial co-localization with lysosomes in the presence of the lysosomal inhibitor leupeptin for US13-V5, US15-V5 and US20-V5. Fibroblasts (HF-TERTs) were mock infected or infected with HCMV Merlin or HCMVs encoding V5-tagged US12 gene viruses at MOI of 10. Lysosomes were stained with LysoTracker (red) at 72 h.p.i. and then fixed. Cells were then co-stained with the anti-V5 antibody (mouse) and anti-mouse Alexa-Fluor AF488 (green), with the nucleus identified with the nuclear stain DAPI. Images were taken using a Zeiss microscope (Axio Observer Z1 with ApoTome) using optional sectioning.

Key: V5 tag= red. Calnexin (ER)= green. DAPI nuclear stain= blue

Figure 5.6

	US14-V5	US15-V5
V5 tag		
Calnexin		
DAPI		
Merge		

Key: V5 tag= red. Calnexin (ER)= green. DAPI nuclear stain= blue

Figure 5.6 cont.

Key: V5 tag= red. Calnexin (ER)= green. DAPI nuclear stain= blue

Figure 5.6 cont.

US18-V5 **US19-V5**

V5 tag

Calnexin

DAPI

Merge

Key: V5 tag= red. Calnexin (ER)= green. DAPI nuclear stain= blue

Figure 5.6 cont.

192

Key: V5 tag= red. Calnexin (ER)= green. DAPI nuclear stain= blue

Figure 5.6 cont.

Key: V5 tag= red. Calnexin (ER)= green. DAPI nuclear stain =blue
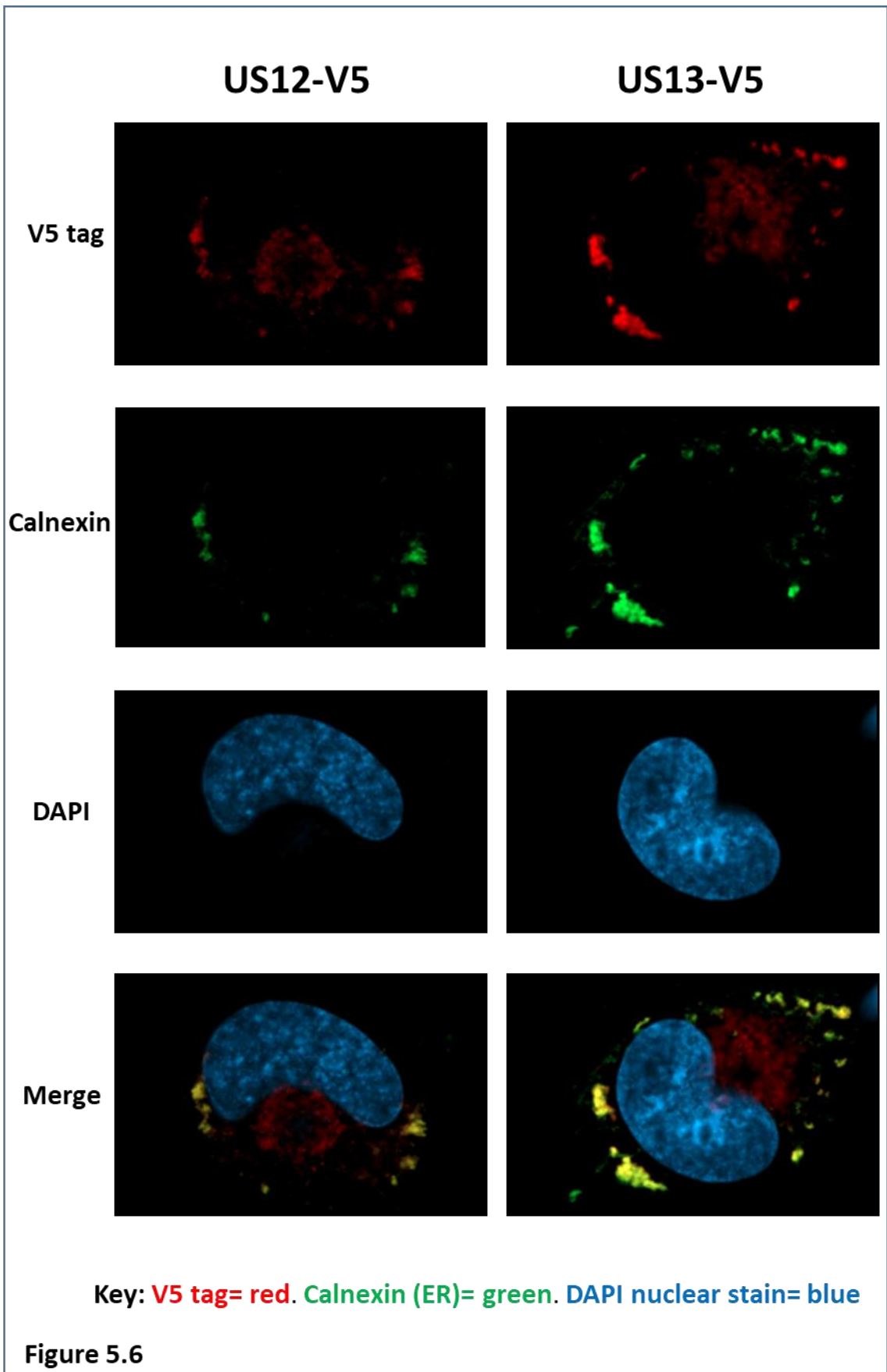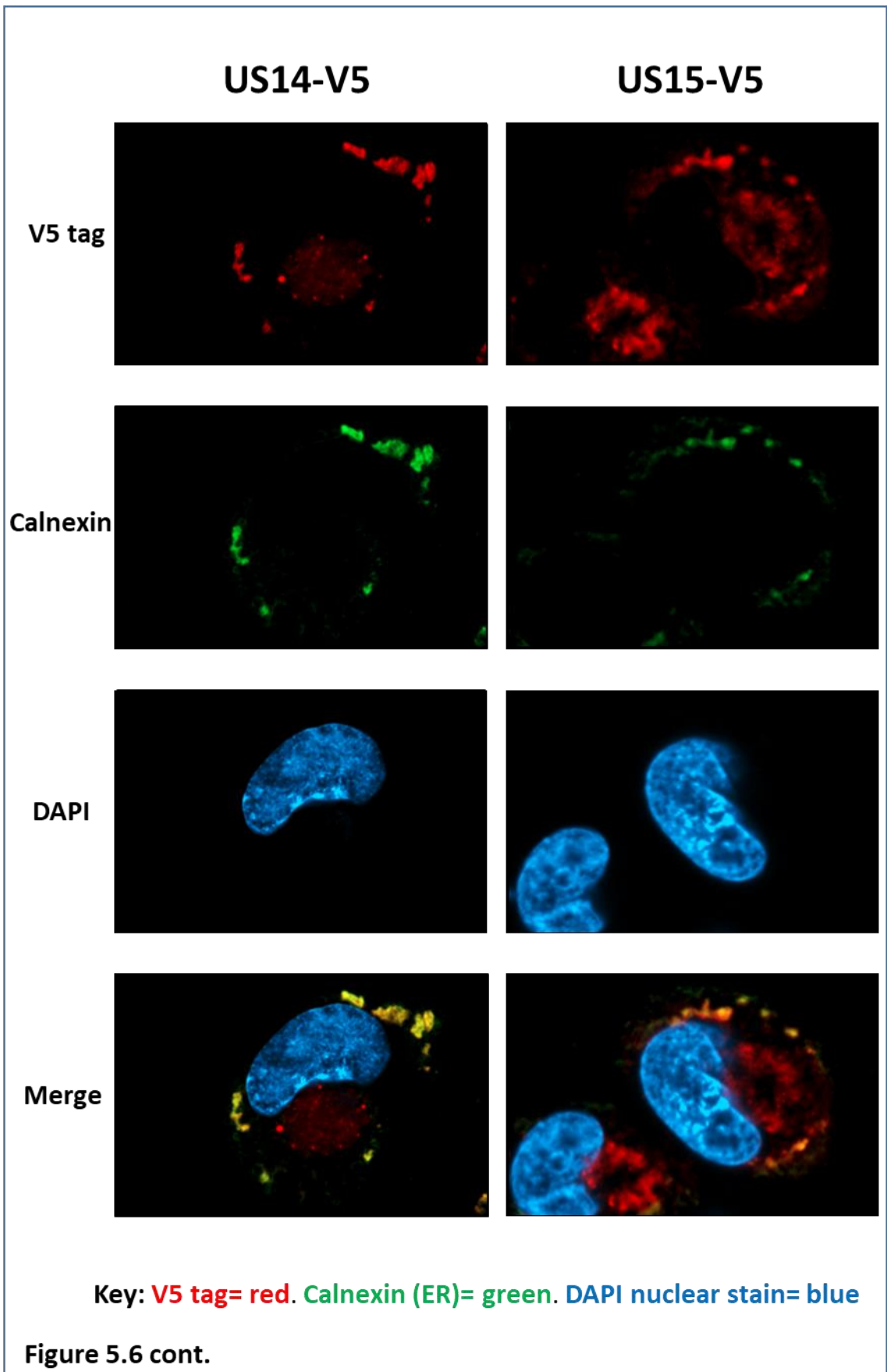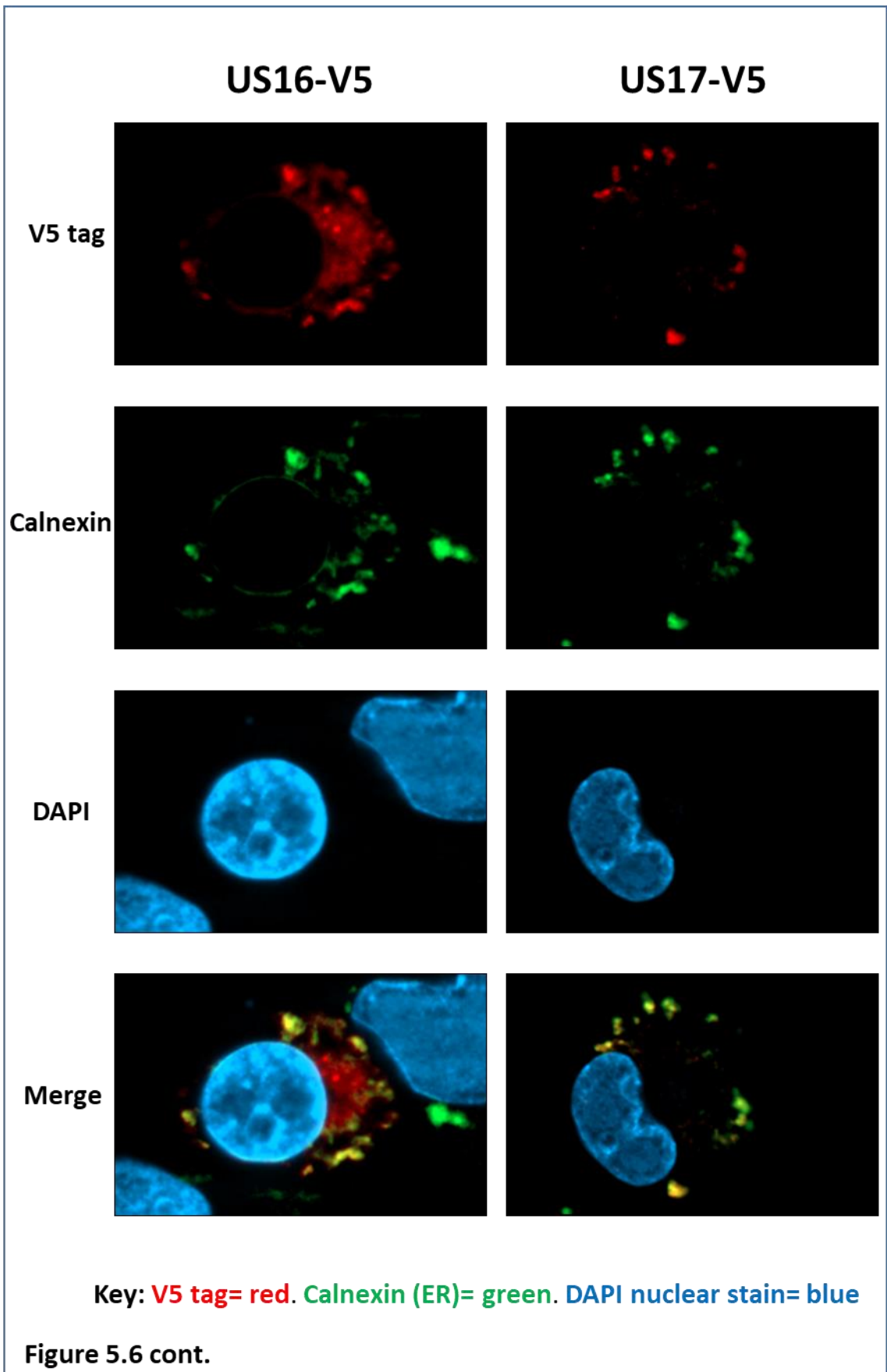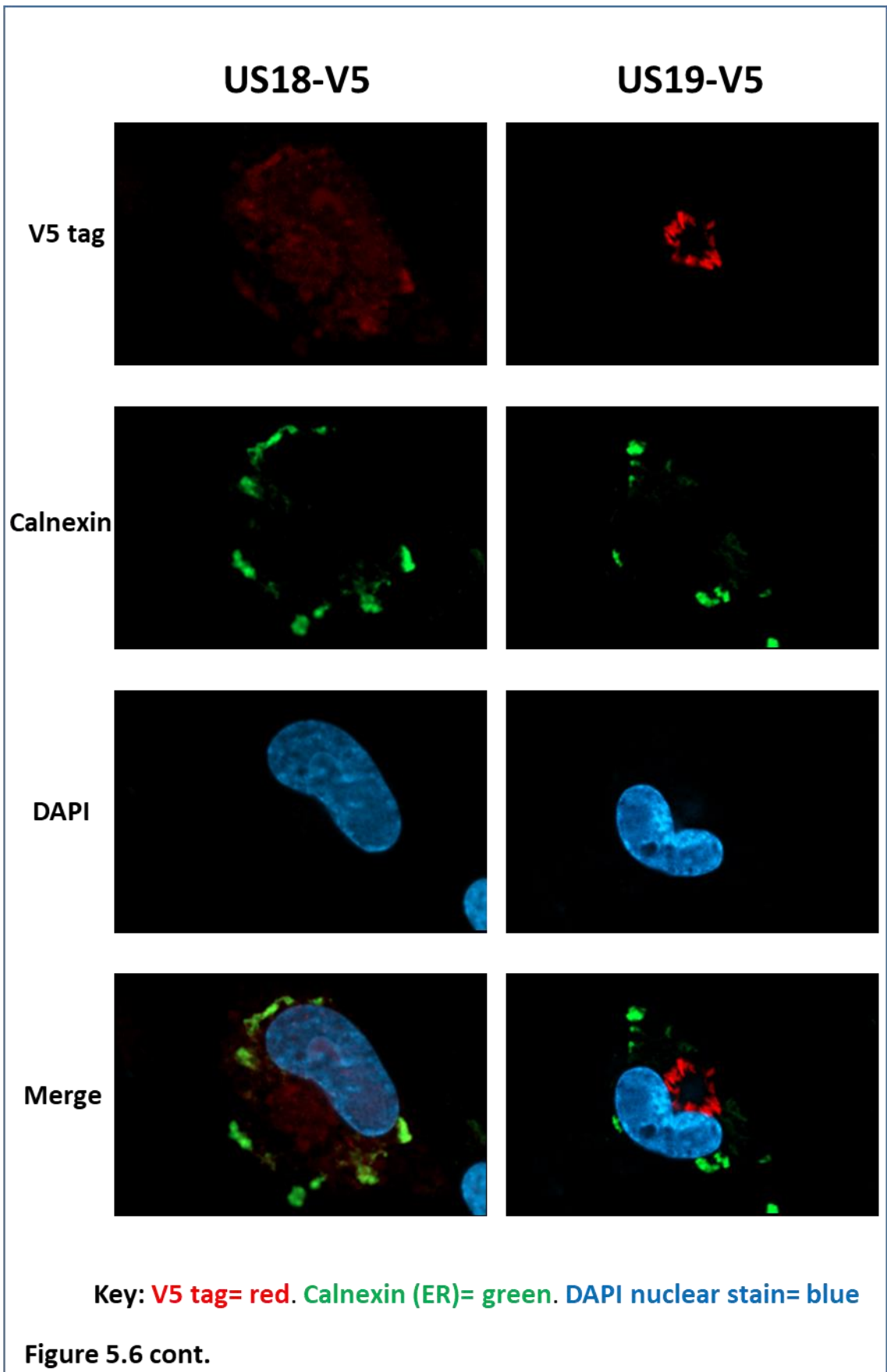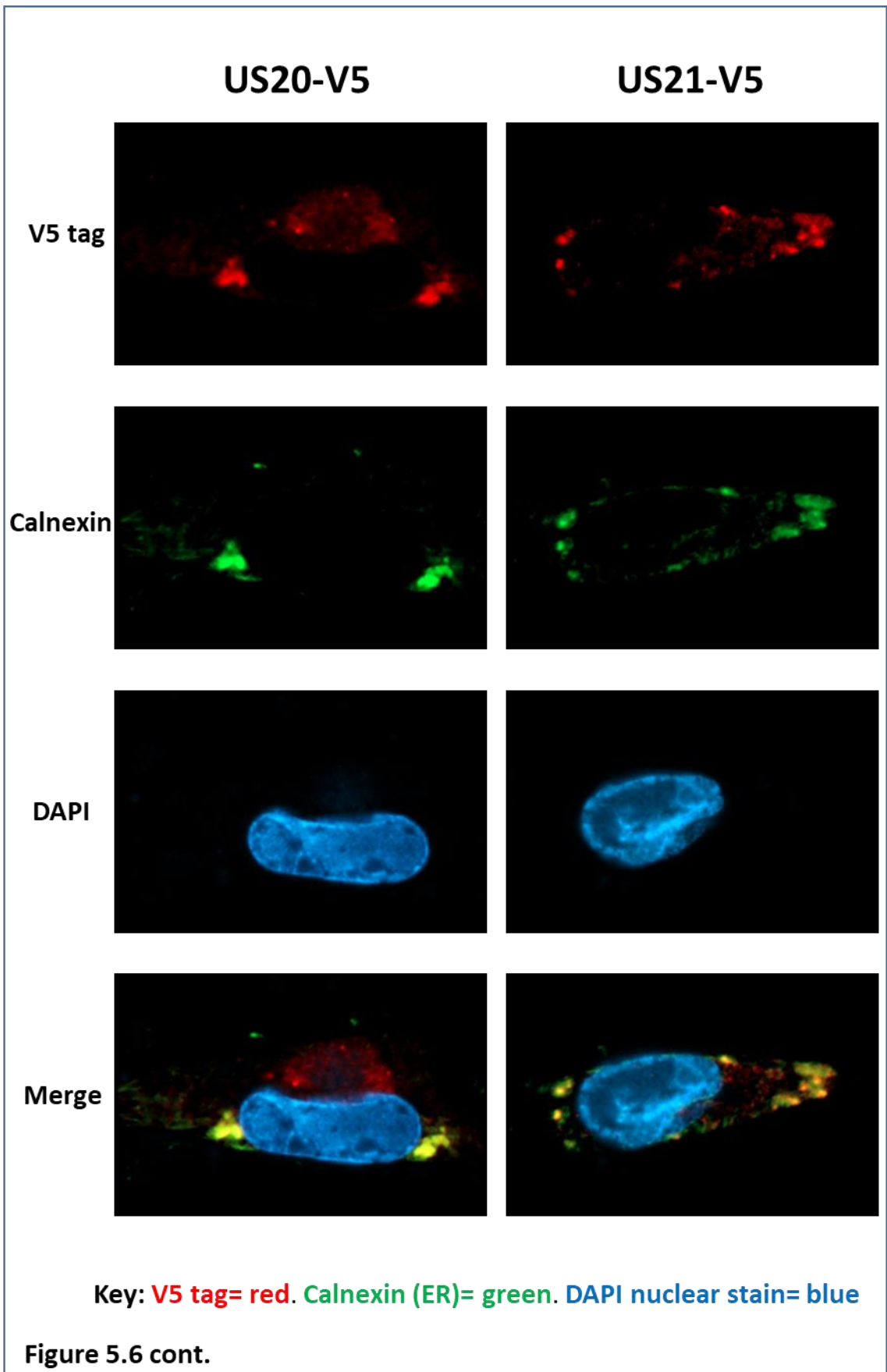
Figure 5.6: Co-localisation of V5-tagged US12 family members with the ER marker Calnexin. Fibroblasts were mock infected or infected with HCMV Merlin or HCMVs encoding V5-tagged US12 gene viruses at an MOI of 10. Cells were fixed at 72 h.p.i with 4% paraformaldehyde. Samples were stained using anti-V5 (rabbit) and anti-Calnexin (mouse) antibodies, with the secondary antibodies anti-mouse AlexaFlour (AF)488 (green) and anti-rabbit AF594 (red), and the nuclear stain DAPI (blue).

The majority of the US12 family members that had ER co-localisation, were also associated with a juxta-nuclear cytoplasmic inclusion (**Fig. 5.6**). This 'dual localisation' was demonstrated for pUS12-V5, pUS13-V5, gpUS14-V5, pUS15-V5, pUS16-V5, gpUS20-V5 and pUS21-V5. In general, the ER co-localisation of these proteins appeared more concentrated, especially for pUS21-V5 in which faint cytoplasmic inclusion localisation could be observed only in some cells (**Fig. 5.6**). gpUS17-V5 on the other hand, co-localised with the ER only, and pUS19-V5 was unique in that it did not localise to the ER but instead exclusively formed around the periphery of the cytoplasmic inclusion.

5.4 Immunofluorescence co-localisation of US12 family members with the virion assembly compartment (vAC)

The juxta-nuclear cytoplasmic inclusion location of multiple US12 family members that was observed alongside their ER co-localisations, was believed to correlate to the virion assembly compartment (vAC). The vAC forms adjacent to the nucleus, causing the kidney shaped nucleus that is characteristic of late HCMV infection, with the vAC forming next to its concave surface (**Section 1.4.1.6**) (Alwine, 2012, Das et al., 2007). In order to determine whether US12 family members associated with the vAC, fibroblasts were infected with the panel of V5-tagged US12 family member HCMV recombinants for 72 hours and the cells stained with an anti-V5 antibody, and an anti-pp28 (ppUL99) antibody as a marker for the vAC (Sanchez et al., 2000b).

The majority of US12 family members were indeed found in association with the vAC (**Fig. 5.7**) although often adjacent to, rather than overlapping, the pp28 marker. This could be partly due to a lower level of protein detection seen for most US12 family members in this location, or their co-localisation with an alternative vAC protein. US16-V5 most commonly showed distinct co-localisation with pp28 and was associated with the outer portion of the vAC, and pUS19-V5 was unique in that it ringed around the periphery of the vAC only (**Fig. 5.7**). gpUS17-V5 was the only member that showed no association with the vAC, and pUS21-V5 only showed occasional faint association, and these members were otherwise observed at the sub-cellular punctate structure (SPS) location that was previously revealed to be the ER (**Section 5.3**). Again, some US12 family members demonstrated 'dual localisation' with both the vAC and SPS defined as the ER. pUS12-V5, gpUS14-V5 and gpUS20-V5 characteristically had dual localisation in the majority of cells examined, whereas pUS13-V5, pUS15-V5 and pUS21-V5 demonstrated occasional dual localisation in a

	US12-V5	US13-V5
V5		
vAC		
DAPI		
Merge		

Key: **V5 tag= red**.
Pp28 (vAC)= green.
DAPI nuclear stain =blue

Figure 5.7

US14-V5 **US15-V5**

V5

vAC

DAPI

Merge

Key: V5 tag= red.
Pp28 (vAC)= green.
DAPI nuclear stain =blue

Figure 5.7 cont.

US16-V5   **US17-V5**

V5

vAC

DAPI

Merge

Key: **V5 tag= red**.
Pp28 (vAC)= green.
DAPI nuclear stain =blue

Figure 5.7 cont.

US18-V5   **US19-V5**

V5

vAC

DAPI

Merge

Key: V5 tag= red.
Pp28 (vAC)= green.
DAPI nuclear stain =blue

Figure 5.7 cont.

US20-V5   **US21-V5**

V5

vAC

DAPI

Merge

Key: V5 tag= red.
Pp28 (vAC)= green.
DAPI nuclear stain =blue

Figure 5.7 cont.

Key: V5 tag= red, Pp28 (vAC)= green, DAPI nuclear stain =blue

Figure 5.7: Some US12 family member proteins are associated with the virion assembly compartment (vAC). Fibroblasts (HF-TERTs) were mock infected or infected with HCMV Merlin or HCMVs encoding V5-tagged US12 gene viruses at MOI of 10. Cells were fixed at 72 h.p.i. and stained for the V5-tagged proteins (anti-V5 rabbit antibodies) and the vAC (anti-pp28 mouse); with Alexa-Fluor AF488 (green) and AF594 (red) as secondary antibodies. The nucleus was identified with the nuclear stain DAPI. Images were taken using a Zeiss microscop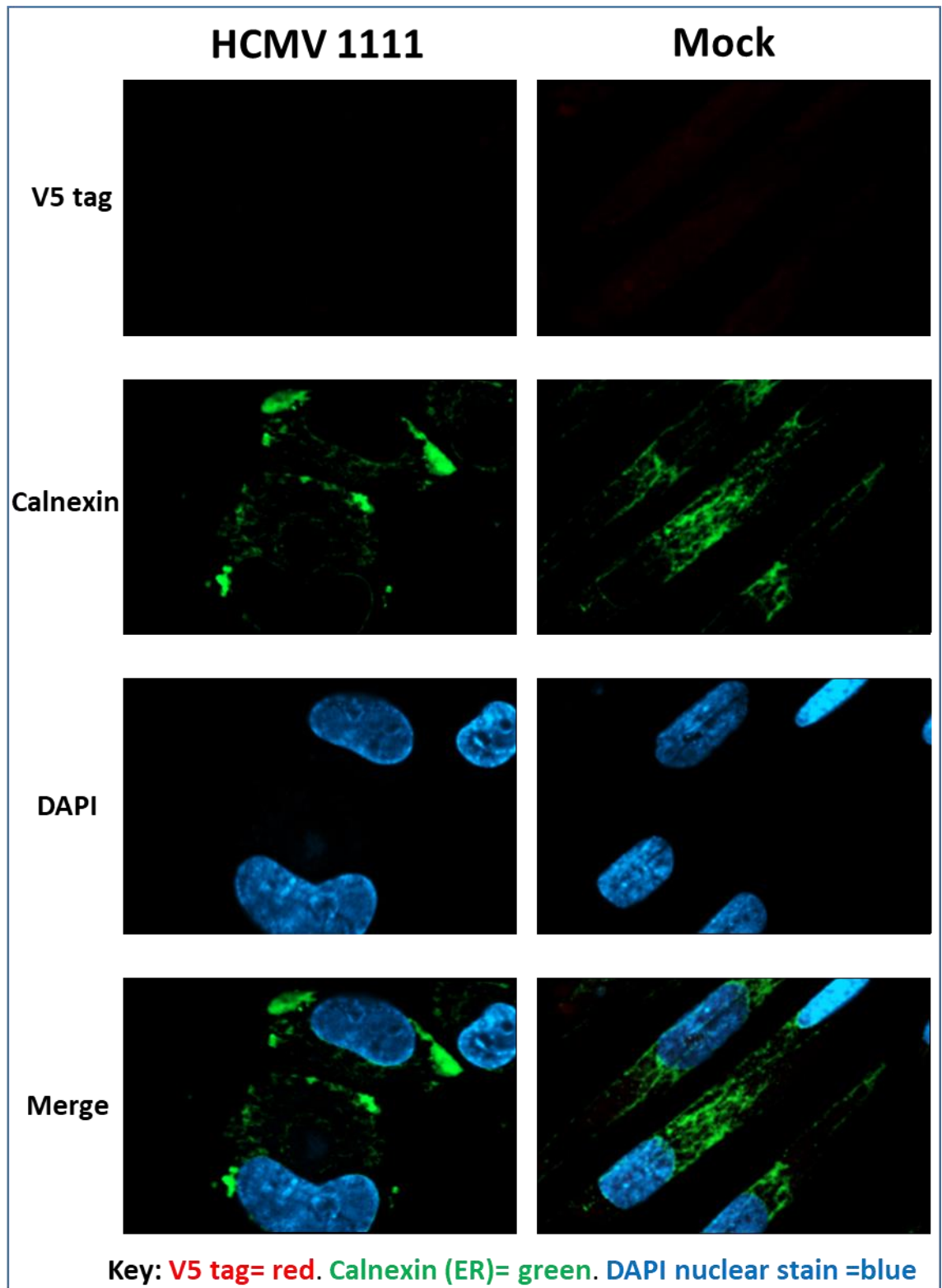e (Axio Observer Z1 with ApoTome) using optional sectioning. Images representative of at least 2 or 3 independent experiments.

subset of cells only. These members generally had less defined vAC associations, and appeared more concentrated at the SPS/ER (**Fig. 5.7**). pUS18-V5 appeared to be disseminated throughout the cytoplasm, and was slightly more concentrated in the vAC of some cells, and occasionally faintly detected at the SPS location.

Although not readily quantifiable by immunofluorescence, US12 family members appeared to be expressed at varying levels of abundance, which corresponded with the findings by immunoblot. Similarly, pUS18-V5 was consistently the hardest member to detect, even when rescued with leupeptin, implying that it perhaps had lower expression levels than the other US12 family members. pUS19-V5 on the other hand was consistently detectable with high efficiency, possibly due to its concentrated localisation around the periphery of the vAC. HCMV control (1111) was used as a control for non-specific staining or potential interactions with HCMV Fc receptors.

In total, 5 members (pUS12, gpUS14, pUS16, pUS19 and gpUS20) showed strong co-localisation with the vAC, 3 members showed weak or occasional vAC co-localisation (pUS13, pUS15, pUS18 and pUS21-V5), and gpUS17 showed no association with the vAC (**Fig. 5.7**).

5.5 Presence of US12 family members in the HCMV virion

Proteins must traffic to the vAC in order to be packaged into virions (**Section 1.4.1.6**). pUS12, pUS13, gpUS14, pUS18, gpUS20 had previously been detected in Merlin virions by proteomics (Murrell, 2014) and these members also exhibited at least partial association with the vAC at 72 hpi (**Section 5.4**). pUS15-V5, pUS16-V5, pUS19-V5 and pUS21-V5 also showed weak or occasional association with the vAC so we wanted to see whether this was consistent with these members also being packaged into virions. In order to test this theory, fibroblasts were infected with the panel of V5-tagged US12 family HCMV viruses and virions were collected by removing the supernatant after the monolayer became 100% infected. In order to reduce sample volume, the viral supernatants were first pumped through vivaflow cassettes to concentrate each sample down to 10-15 ml, and subsequently loaded onto vivaspin columns and centrifuged to reduce the volume further (**Section 2.7**). Both the cassettes and the columns functioned by utilising membranes that allowed water and small molecules to exit, retaining larger macromolecular structures such as HCMV virion particles. Once the sample volume had been reduced to 5-7 ml, they were loaded onto sodium (Na)-tartrate gradients (**Fig. 5.8 A**), 2 tubes per viral sample.

A

Sodium tartrate gradient

'Light'

'Heavy'

B

Viral supernatant

Cellular debris

Virion band

Cellular aggregates

'Heavy' solution (green) = 35% sodium tartrate in NaPh buffer
'Light' solution (blue) = 15% sodium tartrate & 30% glycerol in NaPh buffer
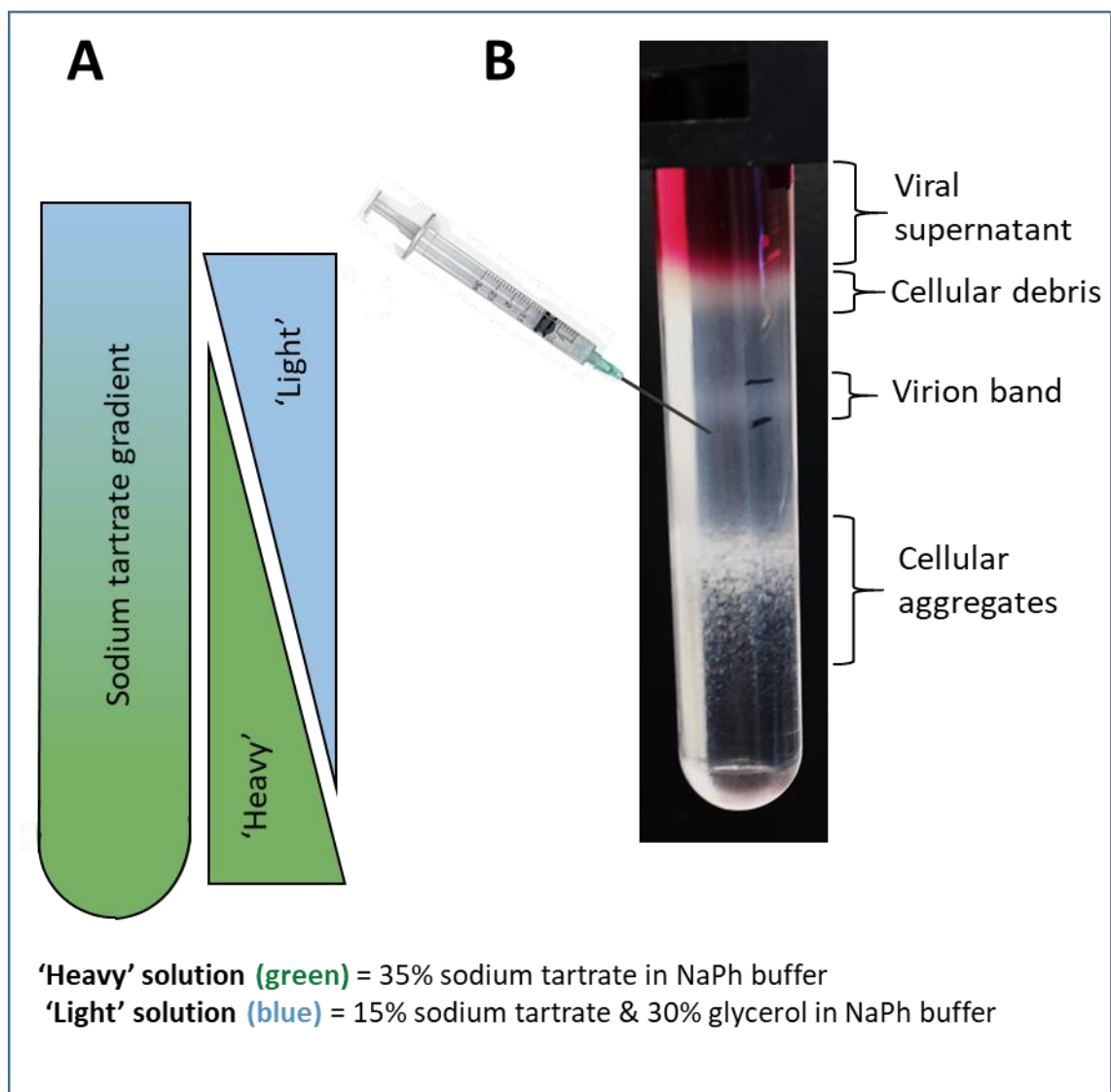
Figure 5.8: Purifying virions using a sodium-tartrate gradient. Virions were purified from viral supernatant collected from fibroblasts (HF-TERTs) infected with HCMV Merlin or HCMVs encoding V5-tagged gene viruses. Supernatants were concentrated using vivaflow cassettes and vivaspin columns down to a volume of 5-7ml. **A)** Sodium tartrate gradients were made by the mixing of 'heavy' (green) and 'light' (blue) gradient buffers, and poured from the top with the use of a gradient maker and a peristaltic pump. Concentrated viral supernatant was loaded onto the gradient, 2 per virus, and centrifuged to separate out the viral components. **B)** An example of the separation seen when using a V5-tagged US12 family member virus on the sodium-tartrate gradient. The virion band forms below the viral supernatant but above the aggregated cellular debris. This band was often faint but could be more easily visualised using laser light. A syringe and needle were used to extract the entire virion layer and after further washing and pelleting steps, the virions could be immunoblotted.

Ultracentrifugation was then undertaken, resulting in the separation of viral and cellular particles over the gradient, with larger molecular weight particles travelling further through the gradient. This separated out the virion particles and cellular debris from the viral supernatant, with the virions forming a distinct band in the gradient (**Fig. 5.8 B**). The virion layer was extracted using a needle and syringe, and the gradient washed off by adding NaPh buffer, and the sample ultracentrifuged to pellet the virions (**Section 2.7**). This pellet was re-suspended in NuPage buffer (diluted in NaPh buffer), and run by SDS-PAGE and immunoblotted. The V5 tagged US12 family members were detected using the V5 tag antibody and were re-probed to measure the level of cross contamination with proteins from the infected cells and supernatant. Calnexin is an abundant cellular protein and along with the virus encoded protein IE1 should not be packaged into the virions, with gB acting as a positive control. Virions were also purified in parallel from characterised HCMV strain Merlin recombinants in which UL148 and UL4 each had C-terminal V5 epitope tags. gpUL148 is a virion glycoprotein that has recently been shown to encode an immunevasin (Murrell, 2014, Wang et al., 2018), and although gpUL4 was historically designated as a virion protein (Chang et al., 1989), more recent studies from our laboratory have revealed that gpUL4 is secreted from infected cells and is not a virion component (Varnum et al., 2004, Murrell, 2014, Seirafian, 2012). As a high abundance secreted glycoprotein, gpUL4 was considered to be the most sensitive control to test for virion contamination, with any proteins of lower or equal expression indicated as unlikely to be in the virion.

US12 family members that were detected in HCMV purified virions included pUS12-V5, pUS13-V5, pUS16-V5, pUS18-V5, pUS19-V5, gpUS20-V5, and pUS21-V5 (**Fig 5.9**). These were split into 2 subsets depending on the level of detection. pUS13-V5 and pUS21-V5 were detected at a relatively low abundance in the virion, with pUS16-V5 only just detectable and was detected barely above the expression levels of the pUL4-V5 negative control. Therefore it could not be confirmed whether pUS16-V5 was a true virion protein or was present due to cellular contamination. Taking into account the lower loading of pUS21-V5 as indicated by gB levels, pUS21-V5 may actually be in the virion in a higher abundance than initially perceived (**Fig. 5.9 A and B**).

The remaining subset of proteins (pUS12-V5, pUS18-V5, pUS19-V5 and gpUS20-V5) were detected as higher abundance virion proteins, with pUS18-V5, pUS19-V5 and gpUS20-V5 consistently detected in relatively large quantities (**Fig. 5.9**). The comparative expression levels of pUS18-V5, pUS19-V5 and gpUS20-V5 correlated with previous immunoblots,
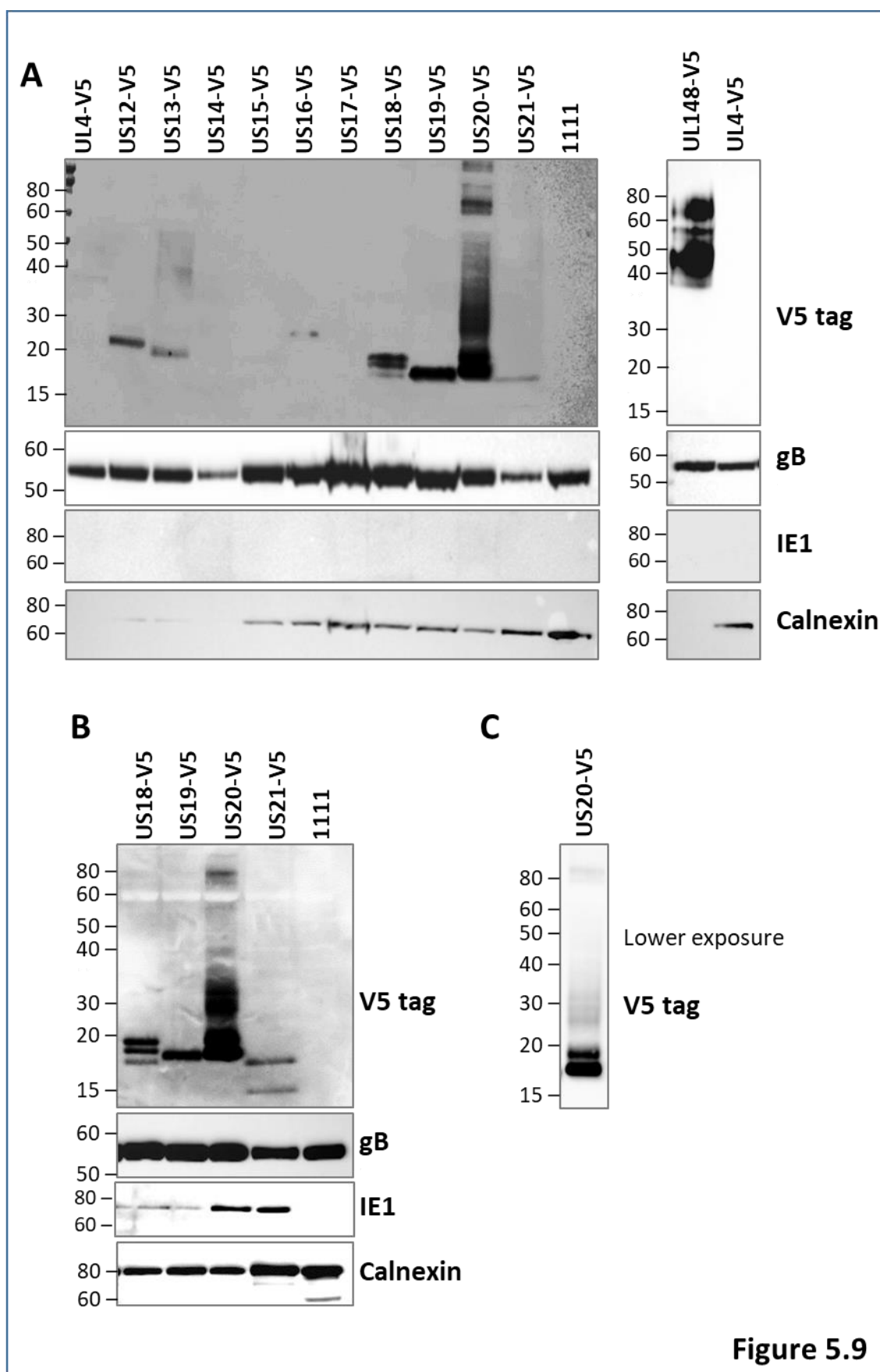
Figure 5.9

Figure 5.9: The presence of the US12 family members in the virion. HF-TERTs in a cell factory were infected with each of the V5-tagged HCMV viruses and 140ml of supernatant was collected ~48hrs after the monolayer was 100% infected. In batches, these were then concentrated using vivaflow cassettes and vivaspin columns and run onto sodium-tartrate gradients and the virion layer extracted (**Fig. 5.4**), washed and pelleted before immunoblotting. The V5 tagged US12 family members were detected using the V5 tag antibody (mouse) and anti-mouse HRP antibody. gB was used as a loading control, IE1 as a negative control and Calnexin as a cellular contamination control. UL4-V5 was used as a 'negative' control due to being a protein suggested not to be in the virion, and UL148-V5 is a known virion protein used as a positive control. **A)** All 10 V5-tagged US12 family member viruses alongside UL4-V5 and 1111 virions, with UL148-V5 virions run on a separate gel. **B)** Samples from A re-run on a new gel to more clearly identify protein bands of pUS18-V5 and pUS21-V5. **C)** Lower exposure image of US20-V5 from A. All samples (except for US16-V5) are representative of 1 independent experiment run on 3 separate blots.

where gpUS20-V5 was the highest expressed and pUS18-V5 was the lowest expressed of the 3 proteins (**Section 3.4**). pUS12-V5 virion expression levels appear to be higher than those of pUS13-V5, although a repeat experiment indicated that they were expressed at more equal levels (data not shown).

Most US12 family members on the virion displayed the same number and same molecular weight of protein bands seen in previous immunoblots (**Chapters 3 and 4**) however the smear/ladder pattern of pUS13-V5 was less obvious in the virion preparation, although a faint detection of a smear could still be seen at ~35-55 kDa. This meant that the reduced detection of the smear was likely due to the lower levels of pUS13-V5 expression seen instead of being due to differential post-translational modifications in the virion. Virion pUS18-V5 was more defined and more readily detected than pUS18-V5 in previous whole cell lysate immunoblots, and an additional faint protein species could be seen at ~18 kDa, situated below the standard doublet band usually detected (visualised here at ~19 and 20 kDa) (**Fig. 5.9 A and B**). This protein species may have been present in previous immunoblots but not detected due to the lower level of expression of this third protein species compared to the doublet, especially considering that WCL pUS18-V5 generally had relatively low expression levels. On the other hand, it is also possible that the virion pUS18-V5 is differentially modified to the cellular pUS18-V5, as previously demonstrated by gpUL141, which presented with an additional protein band in the virion (Cochrane, 2009).

It also appeared that virion pUS21-V5 contained a different number of protein forms to previous WCL immunoblots, with only 1 band appearing to be detected instead of the characteristic doublet. However, upon re-running the sample (**Fig. 5.9 B**), both protein forms were clearly visible and there was therefore no difference between the WCL and virion pUS21-V5. Thus, all US12 family members detected (except potentially pUS18-V5) appeared to be the same size and have the same number of protein species whether found on the virion or in the cell.

The detection of pUS12, pUS13, pUS18-V5 and gpUS20 as virion proteins (**Fig. 5.9**) supports the previous detection of all of these members in the virion by virion proteomics (Murrell, 2014)(**Table 5.1**). The proteomic and immunoblot methods designated pUS13 as a 'low-confidence' or 'low abundance' virion protein, and designated pUS12 and gpUS20 as 'high-confidence' or 'high abundance' virion proteins respectively. Neither method detected pUS15 or gpUS17, with gpUS17 also not detected in the AD169 virion (Gurczynski et al., 2014); and gpUS17 is thus thought not to be present in HCMV virions. pUS19-V5 and

Table 5.1: Detection of US12 family members in the vAC and their presence in virions

Protein	vAC localisation (Fig 5.7)	In the virion by immunoblot? (Fig 5.9)	Merlin Virion proteomics *	Published virion data
pUS12	Yes	Yes	High confidence (7/8 samples)	
pUS13	Yes, occasional	Yes, low abundance	Low confidence (3/8 samples)	
gpUS14	Yes	Not detected	High confidence (8/8 samples)	
pUS15	Yes, occasional	Not detected	Not detected	
pUS16	Yes, peripheral	Yes, low abundance, possible contamination	Not detected	pUS16 not detected in TR virions (Bronzini et al., 2012).
gpUS17	No	Not detected	Not detected	pUS17 not detected in AD169 virions (Gurczynski, Das, & Pellett, 2014).
pUS18	Yes, occasional but faint	Yes, high abundance	Low confidence (1/8 samples)	
pUS19	Yes, rings around the vAC	Yes, high abundance	Not detected	
gpUS20	Yes	Yes high abundance	High confidence (6/8 samples)	
pUS21	**Yes, occasional**	Yes, low abundance	Not detected	

pUS21-V5 were only detected in the virion by immunoblotting and this further supports the observation that immunoblotting is more sensitive than proteomics in detecting US12 family members (**Chapter 3**). gpUS14-V5 was the only member that was detected by proteomics that was not verified by immunoblot as being present on the virion. It is possible that modifications to this protein may have obscured the V5 tag from detection.

5.6 Comparison of US12 family localisations across multiple experiments

The intracellular distribution of US12 family members were generally very consistent within and between experiments (**Table 5.2**). However, there was an observed variation of the distribution of some US12 family members across some cells and experiments. For example, pUS13-V5 and pUS15-V5 co-localised with both the ER and vAC markers in both in the lysosome and ER tracking experiments (**Fig. 5.4 and 5.6**), but in the vAC tracking experiments, they localised mainly to the ER/SPS location, with their vAC localisations less readily observable (**Fig. 5.7).** A similar observation had previously been observed with AD169 pUS14 which was described as being distributed in a uniform granular manner throughout the cytoplasm, concentrating in the AC in some cells only (Das & Pellett, 2007). It seems likely that this 72 hpi time-point is capturing a translocation of these US12 family members between the vAC and ER, causing the slight variations in their observed distributions. Multiple time-points would help to further validate this theory and identify which direction the proteins are travelling in.

The intracellular distribution of US12 family members described above correlates well with the members identified in the Merlin virion, and with previous descriptions in the literature, particularly with studies using the low passage HCMV strain TR (Cavaletto et al., 2015) (**Table 5.1**). Initial studies have suggested that pUS19-V5 may partially co-localise with TGN46, an organelle marker for the trans-Golgi network, although further studies are required (**Figure 5.10**). It was anticipated that US12 family members may co-localise with some of their target proteins, and experiments were set up to attempt to visualise the localisations of target proteins including MICA (1:400 anti-MICA/B BAM01 mouse antibody, Bamomab), B7-H6 (1:200 non-commercial anti-B7-H6 CH31 monoclonal antibody) and PTPRM (1:50 anti-PTPRM 2C10 mouse antibody, Santa Cruz). However, these proteins could not be detected at observable levels within these initial studies, or the proteins appeared not to have a specifically defined location within the cell (data not shown).

Table 5.2: Summary of US12 family localisations from co-localisation studies with lysosomes, the virion assembly compartment (vAC) and the endoplasmic reticulum (ER)

Protein	Lysosome co-localisation? (Fig 5.4)	ER co-localisation? (Fig 5.6)	VAC co-localisation? (Fig 5.7)	Summary
pUS12	No, seen at JCI† and SPS*	Yes, ER plus JCI†	Yes, vAC plus faint SPS*	ER and vAC localisation
pUS13	No, seen at JCI† and SPS*	Yes, ER plus JCI†	Yes, faint vAC plus **SPS***	ER and vAC localisation
gpUS14	No, seen at JCI† and SPS*	Yes, ER plus JCI†	Yes, vAC plus SPS*	ER and VAC localisation
pUS15	No, seen at JCI† and SPS*	Yes, ER plus JCI†	Occasional vAC, plus SPS*	ER and occasional vAC localisation
pUS16	No, but partial crossover due to JCI† location	Yes ER plus JCI†	Yes, outer portion of vAC	Outer portion of vAC and occasional ER localisation
gpUS17	No, seen at SPS* and faint at JCI†	Yes, ER	No, seen at SPS*	ER localisation
pUS18	No, seen cytoplasmic (faint)	Possible faint ER plus JCI†, mainly cytoplasmic	Yes, faint vAC but mainly cytoplasmic	Cytoplasmic dispersive localisation, occasional concentration in vAC and ER
pUS19	No, seen at a JCI†	No, seen to ring around the JCI†	Yes, rings around the vAC	Rings around the vAC periphery
gpUS20	No, seen at JCI† (faint) and SPS*	Yes, ER plus JCI†	Yes, faint vAC plus SPS*	ER and faint vAC localisation
pUS21	No, seen at SPS* and occasional faint JCI†	Yes, ER plus faint JCI†	Occasional vAC, plus SPS**	ER and faint occasional vAC localisation

* SPS= Sub-cellular punctate structure location (separately determined to be the ER)

† JCI= Juxta-nuclear cytoplasmic inclusion location (separately determined to be the vAC)

US19-V5

Key:
V5 tag (green),
Trans-Golgi
Network (TGN46,
red),
DAPI nuclear stain
(blue)

V5 tag

TGN46

DAPI

Merge
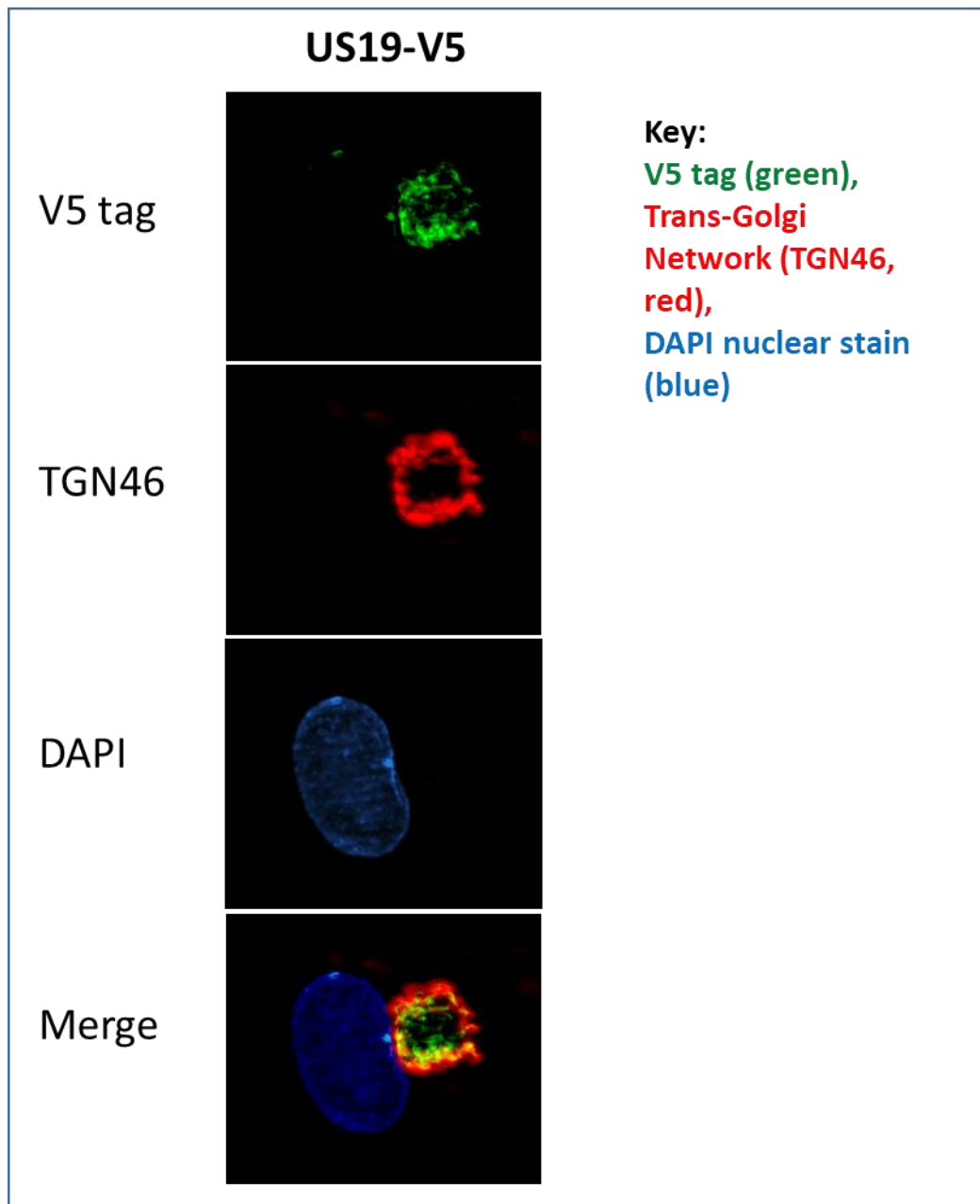
Figure 5.10: The association of US19-V5 with trans-Golgi network marker TGN46. Fibroblasts (HF-TERTs) were infected with HCMV Merlin at MOI of 10. Samples were left untreated, fixed at 72 hpi and stained with an anti-V5 antibody (mouse) and anti-mouse Alexa-Fluor AF594 (red) and co-stained with an anti-TGN46 antibody (rabbit) and anti-rabbit AF488 (green), with the nucleus identified with the nuclear stain DAPI. Images were taken using a Zeiss microscope (Axio Observer Z1 with ApoTome) using optional sectioning. Image is representative of one independent experiment.
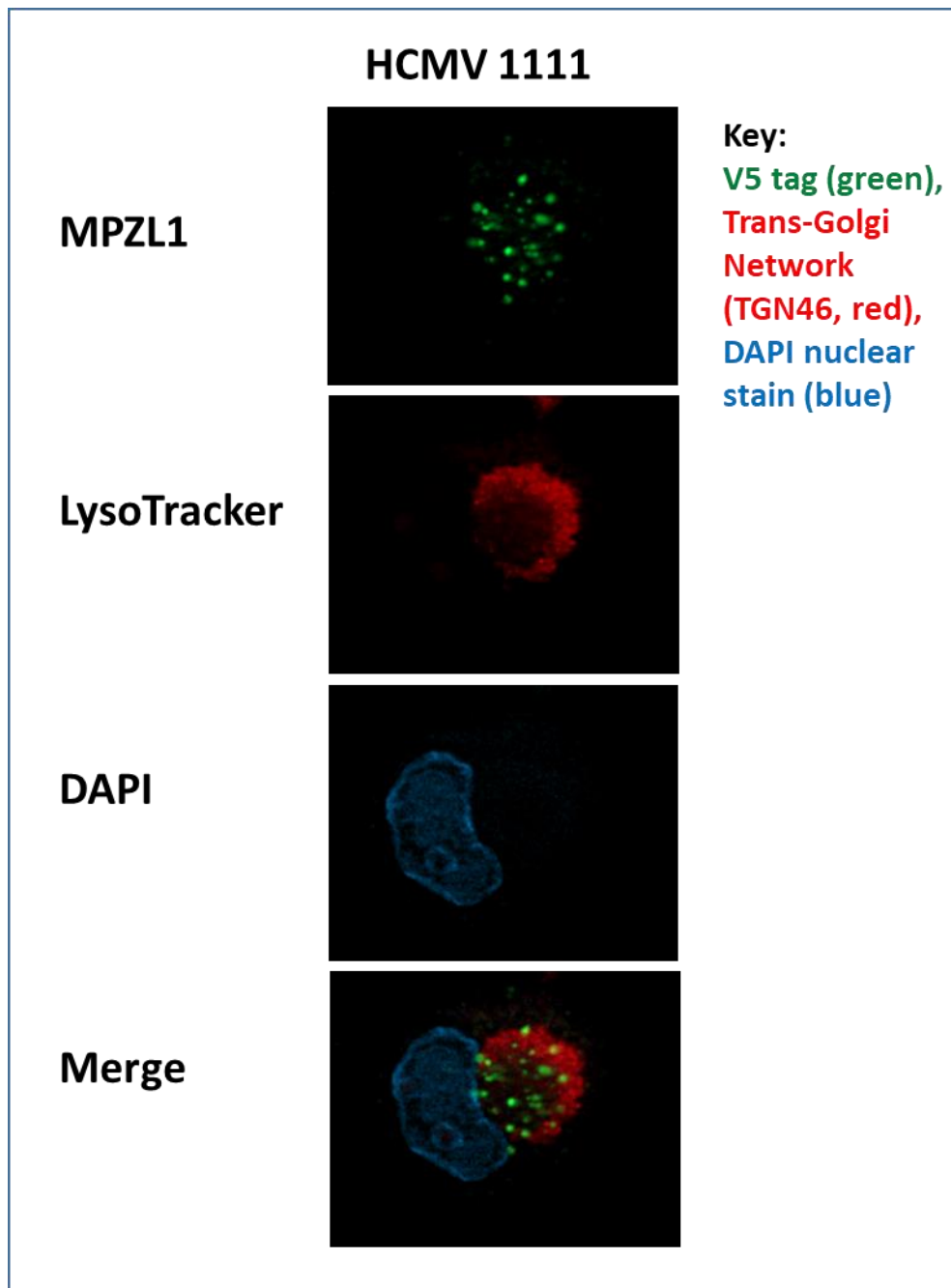
Figure 5.11: The localisation of US12 family target protein MPZL1. Fibroblasts (HF-TERTs) were infected with HCMV Merlin at MOI of 10. Samples were left untreated and lysosomes were stained with LysoTracker (red) at 72 hpi and then fixed. Cells were then co-stained with the anti-MPZL1 antibody (rabbit) and anti-mouse Alexa-Fluor AF488 (green), with the nucleus identified with the nuclear stain DAPI. Images were taken using a Zeiss microscope (Axio Observer Z1 with ApoTome) using optional sectioning.

MPZL1 however was detectable (1:50 anti-MPZL1 H99 rabbit antibody, Santa Cruz) and appeared to concentrate in the vicinity of the vAC, inside the ring of lysosomes, where it had a granular appearance (**Figure 5.11**). Future studies could therefore assess for co-localisation with US12 family members that have been implicated as having MPZL1 as a target protein (Fielding et al., 2017).

5.7 Conclusions

The majority of the US12 family co-localise to some extent in both the ER and the vAC at 72 hpi, with the exception of pUS19-V5 which solely co-localises to the periphery of the vAC, and gpUS17-V5 with solely co-localises to the ER (**Table 5.2**). These findings broadly correlate to the prediction of ER retention motifs in most US12 family members (**Section 4.1.2, Table 4.2**) although other methods of trafficking to the ER are also possible (**Section 6.3.1**). This study provides the first description of the intracellular localisation of pUS12-V5, pUS13-V5, pUS15-V5, pUS19-V5 and pUS21-V5 within HCMV infected cells. Moreover, the association of pUS16-V5 and gpUS20-V5 with the ER and vAC respectively had not previously been demonstrated. These studies also provide evidence that many US12 family members traffic to similar locations within the cell, which would be required if US12 family members act together in complexes (**Section 6.4**).

Seven members of the US12 family were found to be on the HCMV Merlin virion by immunoblot, and this correlated with their vAC localisations by immunofluorescence. Only one member with vAC localisation (gpUS14-V5) was not verified as being incorporated into the virion, although one cannot discount the possibility that its V5 tag could have been cleaved or rendered inaccessible. Access to a gpUS14-specific antibody would help to determine whether this protein was indeed on the virion. pUS21-V5 was the only member detected on the virion that was not regularly co-localised to the vAC. However, an occasional weak association of pUS21-V5 with the vAC in some cells has been revealed, and it would be interesting to investigate whether this association increases at later time points. Virion proteins pUS18 and gpUS20 had previously been shown to have roles in tropism, virion production and content (Hai et al., 2006, Cavaletto et al., 2015) (**Section 1.8.15)**, so their incorporation into the virion is likely to link to these roles (**Section 6.3.2**). pUS16 however could not be confirmed as a virion protein and was also not detected previously in TR virions (Bronzini et al., 2012) (**Table 5.1**). The presence of pUS16 in the virion could therefore be cellular contamination or could be "accidentally" packaged due to its close association with the vAC and its role in tropism.

Immunoblotting combined with an extremely high affinity antibody to the V5 tag has provided a more sensitive method for detecting US12 family members both in the infected cell and purified virions than proteomics (**Chapter 4**, and Weekes et al. (2014)). It is possible that the 7TM domains and extreme hydrophobicity of the US12 family may have presented distinct problems for the proteomic analysis. In the virion proteomics study (Murrell, 2014), even the high abundance US12 family members were not detected in all samples, with pUS20 found in 6 out of 8 samples for example. This may imply that the virus either doesn't always incorporate US12 family members into the virion, or that they may be incorporated in varying amounts. Further immunoblot studies would need to be done in order to validate the frequency of each protein's incorporation into the virion. The demonstration that a subset of US12 family members encode virion proteins agrees and extends published findings and virion proteomics data from our research group (**Table 5.2**) and has provided important new evidence that pUS19 and pUS21 are also virion proteins.

Immunoblotting also supports previous proteomics experiments in relation to the lysosomal degradation of the US12 family members, with both methods detecting the prevention of degradation of pUS12, gpUS14, pUS15, gpUS17, pUS18 and gpUS20 protein expression by leupeptin, a lysosomal inhibitor (this thesis and Fielding et al. (2017)). Immunoblotting additionally indicated that pUS13 was also targeted for degradation and that the degradation of gpUS14-V5 and gpUS17-V5 was detected in their higher molecular weight forms only. This rescue of higher molecular weight forms was also identified in pUS12-V5, gpUS14-V5, pUS15-V5, pUS16-V5 and gpUS20-V5, alongside the rescue of their dominant protein species. These degraded higher molecular weight protein species are likely to be misfolded or incorrectly modified protein forms that were targeted for degradation through the normal cellular degradation process for incorrect proteins. For US12 family members in which the main protein species is rescued from degradation by leupeptin, it is more likely that these proteins are specifically targeted for degradation. As pUS15-V5 is targeted for degradation without containing the TRG_LysEnd_APsAcLL_1 motif, there may be alternative ways of targeting this family for degradation, including through several non-consensus sorting motifs that have previously identified or clathrin-independent routes (Staudt et al., 2017). These increases in protein expression were seen over an 18 h period (54-72 hpi) and further increases would likely be observed over the entire course of infection and could provide better detection and rescue validation of those members which only had minimal rescue in protein expression.

In total, nine US12 family member proteins demonstrated vAC associations and nine US12 demonstrated ER co-localisation within fibroblasts. Despite the rescue of these proteins by the addition of leupeptin, none of the US12 family members appeared to have frequent co-localisation with lysosomes, and the implications of this are discussed further in **Section 6.2**.

6 Discussion

HCMV has the largest genome of any human virus and contains a remarkable array of 15 gene families whose members are generally not essential for virus replication *in vitro* (**Section 1.8**). These gene families together constitute a substantial proportion of the HCMV gene content and many are not shared by other human herpesviruses. Some gene families contain members with similar functions, e.g. both members of the US2 gene family suppress cell surface expression of MHC-1. Therefore, there is a hope that the identification of the function of one member of a family will provide insight into additional members. Many gene families appear to have arisen through a classical 'accordion' gene expansion (Elde et al., 2012), and this can arise after the capture of a host gene, followed by subsequent expansion and divergence. The US12 family is a remarkable tandem array of 10 related genes that appear to have arisen by expansion, possibly from the capture of host G protein-coupled receptor (GPCR) gene (Lesniewski et al., 2006). All US12 family members contain multiple GPCR signature motifs, however none contain the whole set of GPCR motifs and US12 family members have no known GPCR functions (Lesniewski et al., 2006) (**Section 1.8.15**). Phylogenetically, the US12 family represent a distinct branch of the 7TM superfamily, but the closest branches contain transmembrane Bax inhibitor-1 containing motif (TMBIM) family members, with Golgi anti-apoptotic protein (GAAP, CGI-119, TMBIM4) being the nearest human ortholog to any US12 family member (Lesniewski et al., 2006). US21 has the most similarity to TMBIM4 as well as PP1201 (TMBIM1) and LFG (Lifeguard/TMBIM2) (Lesniewski et al., 2006, Holzerlandt et al., 2002) and it was therefore suggested that proto-US21 was the original ancestral gene from which the rest of the family duplicated and diverged (Lesniewski et al., 2006) (**Section 1.8.15**).

The US12 family members are highly conserved, despite being dispensable for replication *in vitro* and despite HCMV displaying the highest level of genetic diversity of all human herpesviruses (Dunn et al., 2003b, Sijmons et al., 2015)(**Section 1.8.15**), indicating that their functions are important for HCMV persistence *in vivo*. US18 and US20 were the first US12 family members to be identified as having NK evasion functions (Fielding et al., 2014), which they achieve by controlling the cell surface levels of MICA (**Section 1.8.15.4**). It was hypothesised that the whole family could be affecting the cell surface proteome, so a proteomics approach was undertaken. The US12 family were demonstrated to regulate the cell surface proteome, and their cellular targets were identified (Fielding et al., 2017). This thesis aimed to complement this study, by further characterising the US12 family gene products for their expression, degradation and localisation patterns. A fundamental

appreciation of US12 family gene expression was important in order to understand the proteomic data.

HCMV proteins are often not processed correctly or function in the same way outside of the context of infection, and previously, experimental data had indicated that the US12 family members also demonstrated a lack of full function when expressed individually from adenovirus vectors (Dr C. Fielding, as detailed in **Section 3**), and this thesis has further determined that the US12 family had altered expression levels and different N-glycosylation patterns when studied in isolation, and some members had limited protein detection by immunoblot when the sample was untreated (**Figure 4.4**). The difficulties of studying US12 family members in isolation is likely to be due to the lack of other HCMV proteins required for the US12 family to be expressed, processed or to function correctly. Without the context of infection, there is also a lack of rearrangement of the host cellular machinery to form the viral assembly compartment (vAC), which may distort the US12 family protein's localisation and may affect interactions with cellular and viral binding partners. Consequently, this thesis studied the US12 family by inserting a V5 tag at the C-terminus of each US12 family member within the HCMV Merlin BAC genome.

6.1 Characterisation of HCMV US12 family expression and post-translational modifications.

Three US12 family members were identified as containing N-glycosylation within the context of infection: gpUS14-V5, gpUS17-V5 and gpUS20-V5 (**Section 4.4**). gpUS17-V5 and gpUS20-V5 were detected as a doublet by immunoblot, with only the higher molecular weight species sensitive to digestion with endoglycosidase H (EndoH). N-glycosylation can be important for protein folding, stability and function, with many glycosylated proteins found in the HCMV virion envelope, however only gpUS20-V5 was identified in the Merlin virion by immunoblot (**Section 5.5**), although gpUS14-V5 had been detected in the virion previously by proteomics (Murrell, 2014).

This finding also corresponds with the prediction that the C-termini of these three US12 family members would be non-cytoplasmic, allowing all or the majority of their N-glycosylation sites to be facing the lumen of the ER and accessible to the glycosylation machinery (**Section 4.1.3**). This also indicates that the US12 family members are likely to have opposing membrane topology to 7TM GPCRs, which generally have non-cytoplasmic

N-termini (Lesniewski et al., 2006), however TMBIMs generally have cytoplasmic N-termini (**Figure 6.1**). For some TMBIMs, such as GAAP/TMBIM4, it has recently been discovered that their last transmembrane domains ('TM7') were not true transmembrane domains, but instead formed probable re-entrant loops, resulting in both the C- and N-termini belonging to the cytosolic side of the membrane (Carrara et al., 2012) (**Figure 6.1**). Although some US12 family members were predicted to contain re-entrant loops themselves by certain prediction software programs (**Section 4.1.3**), no re-entrant loops were predicted for any of the US12 family's TM7 domains. This is likely because the US12 family have more hydrophobic TM7 domains than those of TMBIM4 and TMBIM6, as determined for gpUS20 (Cavaletto et al., 2015). Some US12 family members were also predicted to contain only 6 transmembrane domains, however the overall consensus remains that each US12 family member most likely has 7 TMDs, and gpUS20 was experimentally demonstrated to have a cytosolic N-termini and its C-termini in the ER lumen (Cavaletto et al., 2015). The structure of GPCRs and TMBIMs appears to be important for their function so the US12 family topology being the reverse of GPCRs suggests that they may have a non-GPCR related function. As they do appear however to have their N-termini on the same side of the membrane as TMBIM members, of which they are more closely related (**Figure 6.1**)**,** it does not rule out TMBIM-related functions being possible for US12 family members.

Aside from N-glycosylation, there appears to be other post-translational modifications (PTMs) of US12 family members, with many members demonstrating protein forms of different molecular weights or ladder/smear patterns by immunoblot (**Section 3.4, 4 and 5.1**). A smeared pattern can be caused by the nature of running hydrophobic transmembrane proteins by SDS-PAGE, as they can often appear to aggregate at high temperatures or may not fully denature, and this had previously been noted for pUS18 and pUS20 (Fielding et al., 2014). However, ubiquitination and SUMOylation of proteins can also result in a laddering or smear pattern (Choo and Zhang, 2009, Seyfried et al., 2008, Wang et al., 2017). pUS13-V5 demonstrates the most extensive ladder pattern across immunoblots, with pUS12-V5, gpUS14-V5, pUS15-V5, pUS16-V5, gpUS17-V5, gpUS20-V5 and pUS21-V5 demonstrating lesser degrees of smears or ladders, often only detected at longer exposure times or upon the addition of leupeptin (**Chapter 3.4, 4.2, 4.4 and 5.1**). If the ladder and smear patterns of US12 family members is due to ubiquitination, this could mean that this subset of US12 family proteins could be targeted for ubiquitin-mediated degradation in the proteasome. Ubiquitin can also affect cellular
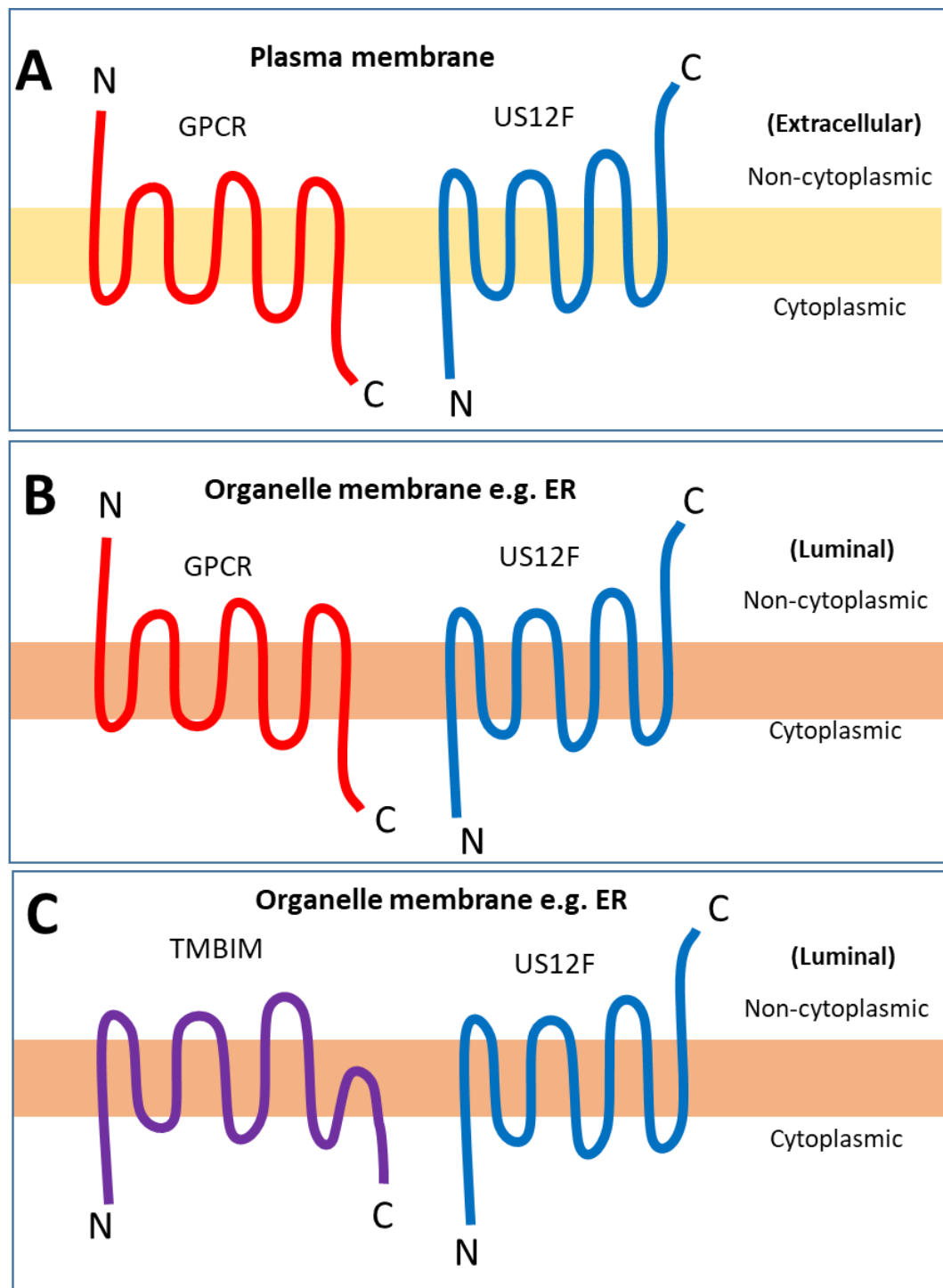
Figure 6.1: Membrane topology predictions for the US12 family, GPCRs and TMBIMs. Predicted or proven membrane topologies **A)** GPCR and US12 family topologies across the plasma membrane. **B)** GPCR and US12 family topologies across an organelle membrane. **C)** TMBIM and US12 family topologies across an organelle membrane, based on the findings of TMBIM4 and TMBIM6 (**Section 6.1**).

localisations, activity and can promote or prevent protein interactions (Schnell and Hicke, 2003). This possibility could be established by immunoprecipitating each V5-tagged protein and immunoblotting the sample with an anti-ubiquitin antibody. SUMOylation is analogous to ubiquitination and can change the localisation of proteins, their stability or enzymatic activity, and their interactions with cellular components or binding partners (Boggio and Chiocca, 2006). Viral proteins can also influence SUMOylation, and some viral proteins need to be SUMOylated in order to function correctly (Boggio and Chiocca, 2006). SUMOylation of US12 family members could be tested through the use of a cell line that expresses His-tagged SUMO available at Cardiff (Prof. G. Wilkinson, personal communication). There currently seems to be no correlation between those US12 family members with ladder or smear patterns and those without in terms of localisation, and no link can be made to protein function, with these US12 family members having a range of functions, and the most extensively laddered protein pUS13-V5 having no known function.

Expression of all members of the US12 family was tracked using a high affinity antibody to the V5 epitope tag. If the V5 tags were always recognised with equal efficiency, then the different US12 family members are demonstrated to be expressed at different levels to each other. These differences could be quite extreme, with gpUS20-V5 detected at 33 times higher levels than pUS21-V5 by densitometry (**Section 3.4**). Three distinct 3' co-terminal transcripts have been reported for the US12 family, that encode for US21 alone, US20-US18 and US17-US12 (Guo, 1993, Towler, 2007, Gatherer et al., 2011). gpUS20-V5 was expressed more efficiently than pUS19 and pUS18 in the context of HCMV infection (**Section 3.4**), consistent with it being with the first ORF on the polycistronic transcript (Guo, 1993). pUS18 was detected at 20 times lower levels than gpUS20-V5 by densitometry and consistently had the lowest expression of the 3 proteins (**Section 3.4**), potentially because ribosomes can be expected to translate downstream ORFS with reduced efficiency due to falling off prematurely. Although an attractive model, the US17-US12 transcript set did not appear to follow such a pattern of expression (**Section 3.4**). Therefore, lower expression of US12 family members may also have been the result of quicker protein turnover or an obstruction to the detection of the V5 tag due to protein folding or similar. It has also been suggested that there are additional transcripts within the US12-US17 region of Han and AD169 HCMV strains that may subsequently affect the protein levels of each US12-US17 member **(Fig. 6.2)** (Lu et al., 2016)**.** US21 is individually situated on its own transcript (Guo, 1993) so may have its own transcriptional and translational control, and in
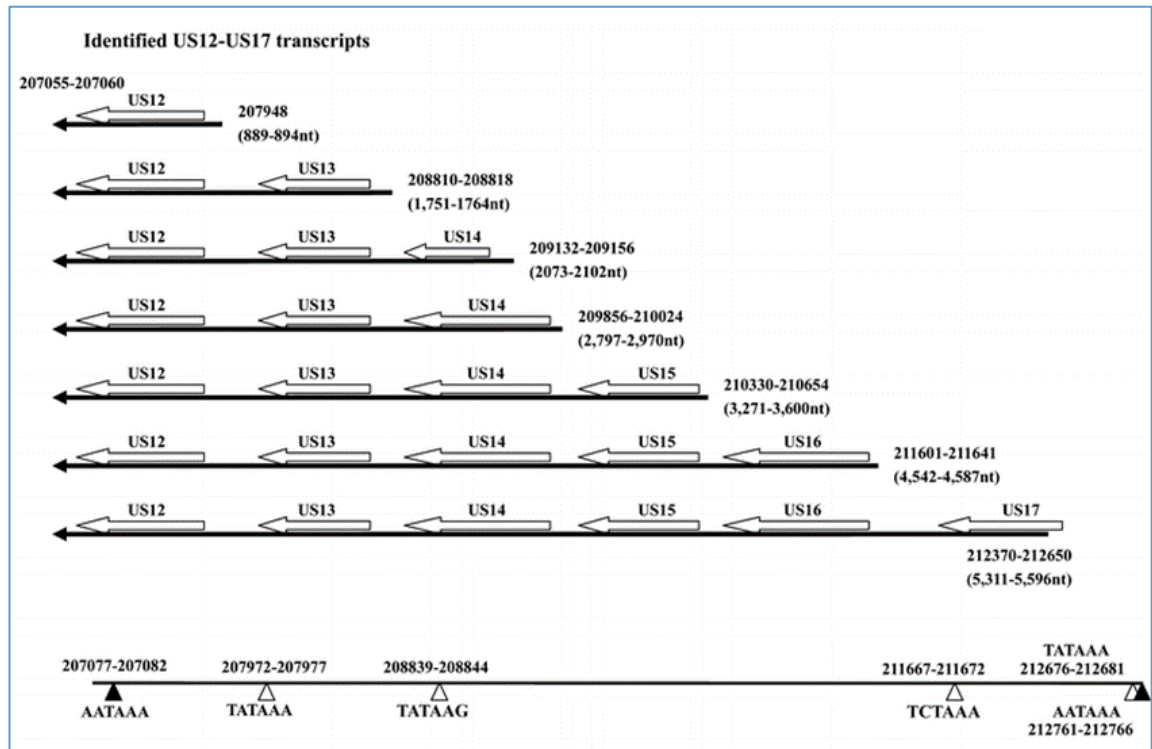
 Figure 6.2: Transcriptional map of US12-US17 from HCMV Han (Lu et al., 2016). The predicted ORFs (indicated by hollow arrows) of HCMV Han in the US12-US17 region, as taken from Lu et al. (2016), figure 7. TATA elements (indicated by hollow triangles) and polyA signals (indicated by black triangles) are indicated at the bottom of the figure. Transcripts detected have their 5' and 3' ends of the transcripts labelled, with their approximate lengths given in brackets, as estimated by Northern Blot and RACE.

general, has lower expression levels than most other US12 family members (**Section 3.4**). It cannot be ruled out that the addition of the V5 tags could have affected the transcriptional levels of the US12 family in some way, which may have in turn affect the protein levels. Although the V5 tags of the US12 family do not appear to overlap with any known TATA elements or polyA signals, it would be beneficial to study the transcript levels with and without V5 tagging to provide evidence for or against this possibility. It is also possible that the V5 tag may have affected the stability or degradation of US12 family members (**Section 4.6**).

At least five of the US12 family members have their main protein form targeted for degradation within the context of HCMV infection which would also affect the amount of protein available for detection, with pUS12-V5 demonstrating up to a 15 fold increase upon the addition of Leupeptin by densitometry (**Section 5.1**)**.** The detection of pUS13-V5 as a smear/ladder may also partly explain why this protein was harder to detect than if the total protein amount was found in a single molecular weight band. This is supported by densitometry that demonstrated that in some experiments the pUS13-V5 sample contained as much protein as the gpUS20-V5 sample but appeared much weaker due to the ladder effect (**Section 3.4**). Outside of the context of infection, these expression patterns weren't followed, suggesting that these differences in expression were related to HCMV infection, however Ad-US14-V5 and Ad-US17-V5 had previously been codon optimised which had increased their protein levels (**Section 4.4**).

The nature of the protein expressed by each US12 family member remained constant over the course of a productive HCMV infection and where relevant, was generally similar to the protein packaged in to the virion with respect to its molecular weight and banding pattern by immunoblot (**Section 4.2 and 5.5**). pUS18-V5 was the only family member in which an additional protein band could be detected in the virion, and this may be due to alternative processing or modifications to the virion form of the protein. It is possible that this additional pUS18-V5 form is N-glycosylated, as other HCMV proteins such as gpUL141 have shown to be differently glycosylated in the virion than in the WCL (Cochrane, 2009). An attempt to verify this was made, but not enough protein was present for detection of pUS18-V5 after de-glycosylation treatment. Glycosylation generally increases the molecular weight of a protein however, whereas this additional protein form of pUS18-V5 had a lower molecular weight (18 kDa) than the standard doublet detected (**Section 5.5**). The simplest explanation was that this band could not be detected in previous immunoblots due to its

lower abundance, as it was detected at a lower expression level than the original doublet, and because pUS18-V5 consistently had low expression levels in the WCL. There is also the possibility that this lower band could relate to the truncated US18 transcript that was detected in Merlin from 72 hpi (Stern-Ginossar et al., 2012) (**Section 4.6**) which could correlate with the incorporation of this 18 kDa pUS18-V5 protein into the virion.

Immune evasion proteins are often required early in infection to prevent the host mounting an immune response to HCMV. For example, US3 is expressed as a Tp1 (immediate early) protein and functions early to prevent antigen processing and presentation at early stages of infection in order to limit immune recognition (Colberg-Poley, 1996, Noriega et al., 2012b). Therefore it was expected that the US12 family would have relatively early expression patterns if involved in immune evasion, as US18 and US20 had demonstrated to be (Fielding et al., 2014) and as indicated by the US12 family's ability to regulate cell surface immune ligands (Fielding et al., 2017). Instead, the US12 family mostly belonged to the Tp3 temporal expression kinetic class (as described in **Section 1.4.1.4**) but ranged from Tp2 early proteins (pUS16-V5 and gpUS20-V5) to Tp5 late proteins (gpUS17-V5), with the majority of members having their highest expression levels at 72 hpi. Although immediate early time-points were not collected, all family members could be detected by either 24 hpi (pUS12-V5, gpUS14-V5, pUS16-V5, pUS18-V5 and pUS20-V5) or 48 hpi (pUS13-V5, pUS15-V5, gpUS17-V5, pUS19-V5 and pUS21-V5) (**Section 4.2**) and members may be expressed from earlier time-points, even if they are not abundant at those times. Their broad expression patterns might correlate with their modulation of cellular immune ligands including MICA and MICB which accumulate in the cell at high levels at multiple time-points without the presence of the US12 family members (Fielding et al., 2017). The requirement of different US12 family members at different times may also indicate towards different functions. pUS16-V5 and gpUS20-V5 had high expression levels across all time-points so could have a role that either requires their expression throughout the duration of infection or could have both early and late roles. pUS16-V5 and gpUS20-V5 were also the only Tp2 expressed proteins, and along with pUS12-V5 were the only US12 family members to have their highest relative expressions at 48 hpi. pUS12 and gpUS20 have been demonstrated to play a role in NK evasion, along with gpUS14 and pUS18 which are Tp3 proteins (Fielding et al., 2014, Fielding et al., 2017) and this role would likely be required from early time points, with all 4 of these members detected by 24 hpi (**Section 4.2**). Their NK evasion roles also link with the known targets of these US12

family members, with US20 able to target the NK ligands MICA, MICB, ULBP2 and B7-H6 (Fielding et al., 2017, Fielding et al., 2014). US18 was also able to target MICA and B7-H6 in concert with US20, and US12 was able to target ULBP2 (Fielding et al., 2014, Fielding et al., 2017). There are currently no known NK ligand targets for pUS14 however. One of the target proteins of the US12 family is pUL16, an NK evasion protein that functions by sequestering NK ligands such as MICB, ULBP1, 2, 4 and 6 in the ER (Rolle et al., 2003). Deletion of US12, US13 or US20 lead to increased cell surface and intracellular expression of UL16, and the deletion of US16, US17 or US18 appeared to lead to an increase of UL16 in the plasma membrane only (Fielding et al., 2017). This may show functional redundancy of NK ligand targeting between UL16 and the US12 family, or it may be that UL16 becomes degraded indirectly through its interaction with MICB and ULBP2 which are targeted for degradation by the US12 family.

Being categorised as Tp3 does not diminish their capacity to be involved in immune evasion, with other proteins classified as Tp3 (Weekes et al., 2014) demonstrated to be involved in immune evasion, nuclear egress and protein translocation, including UL40 (Prod'homme et al., 2012), UL50 (Sharma, Kamil, Coughlin, Reim, & Coen, 2014), IRS1 (Child, Hakki, De Niro, & Geballe, 2004; Ziehr, Vincent, & Moorman, 2016) and US6 (Lehner, Karttunen, Wilkinson, & Cresswell, 1997). Even the Tp5 late protein gpUL141 targets NK ligands (CD112 & CD155) (Prod'homme et al., 2010, Tomasec et al., 2005, Weekes et al., 2014), so the possibility that pUS21 (a Tp5 protein) also contributes to NK evasion (Fielding et al., 2017) is still plausible.

Non-immune evasion functions may require expression of the US12 family at later time-points, with both pUS16 and gpUS20 having published roles in tropism which would be required later in infection (Bronzini et al., 2012, Cavaletto et al., 2015, Luganini et al., 2017). pUS16 alters tropism by reducing the content of the pentamer on the virion, which prevents entry into endothelial and epithelial cells (Bronzini et al., 2012, Luganini et al., 2017). Altering the virion content or its maturation can not only have tropism effects, but can also cause early effects on the next newly infected cell. For example, the deletion of US17 causes a larger amount of non-infectious viral particles, which results in a 2.6-fold increase in the level of intracellular pp65 delivered to each cell, which is the likely cause of the markedly blunted the host cell antiviral response in the newly infected cell (Gurczynski et al., 2014). US17 can both differentially regulating transcripts at 96 hpi compared to the AD169 parental virus, as well as differentially regulate transcripts at 12 hpi in the recently

infected cell, effectively having both early and late functional effects, despite being a late protein. Proteins that are incorporated into the virion themselves would also be expected to be expressed later in infection. pUS21-V5 for example is a Tp5 protein and has been shown to be present in HCMV Merlin virions (**Section 4.2 and 5.5**) although US21 currently has no known target proteins or functions to explain its late expression pattern, or why its presence in the virion may be required.

6.2 Regulation and degradation of US12 family members and their targets

The US12 family target cellular proteins and direct a subset of them for degradation (Fielding et al., 2014, Fielding et al., 2017) (discussed in **Section 1.8.15.3**), and their ER localisations may allow them access to these proteins to achieve this. The cell surface immune ligands that are targeted by the US12 family would have to pass through the ER and Golgi to arrive at the plasma membrane. The US12 family members that localise to the ER would thus would be in a position to interact with their target proteins and could direct these proteins for degradation before they were translocated to the cell surface. It was hypothesised that US12 family members may directly transport their target proteins for degradation and that the degradation of US12 family members (**Section 5.1**) may be an indirect consequence of this. In support of this theory, many of the US12 family (and their targets) could be rescued by leupeptin, generally recognised to be a lysosomal inhibitor. US12 and US13 had also previously been identified as being at least partially belonging to the lysosomal fraction of HCMV AD169 organelles, with US12 identified at 24 and 48 hpi and US13 at 120 hpi (Jean Beltran et al., 2016).

The lysosomal targeting of cellular proteins is also a property of the TMBIM family, with TMBIM1 and 3 targeting TLR4 and Gb3 synthase for lysosomal degradation respectively (Yamaji et al., 2010, Zhao et al., 2017). It is also possible that instead of directing their target ligands for degradation from the ER, they could be targeting them from the plasma membrane (PM) instead. Many US12 family members were detected in the PM by proteomics (Weekes et al., 2014, Fielding et al., 2017), and although they have not been identified on the PM by immunofluorescence, this could be because their abundance on the PM was too low or too transient to detect. If the US12 family do translocate to the PM, they would likely be retrieved by the endosomal transport system, in a similar way to how virion glycoproteins are retrieved (**Section 1.4.1.7, Figure 6.3**). These endosomal vesicles

Figure 6.3

Plasma membrane

6. Virions transported to cell surface

Could endosomes fuse with lysosomes to cause the degradation?

4. Retrieved by endocytic system

5. Transported to vAC and used to envelope capsids

3. Exported to PM

vAC

2. Processed through Golgi and TGN

ER
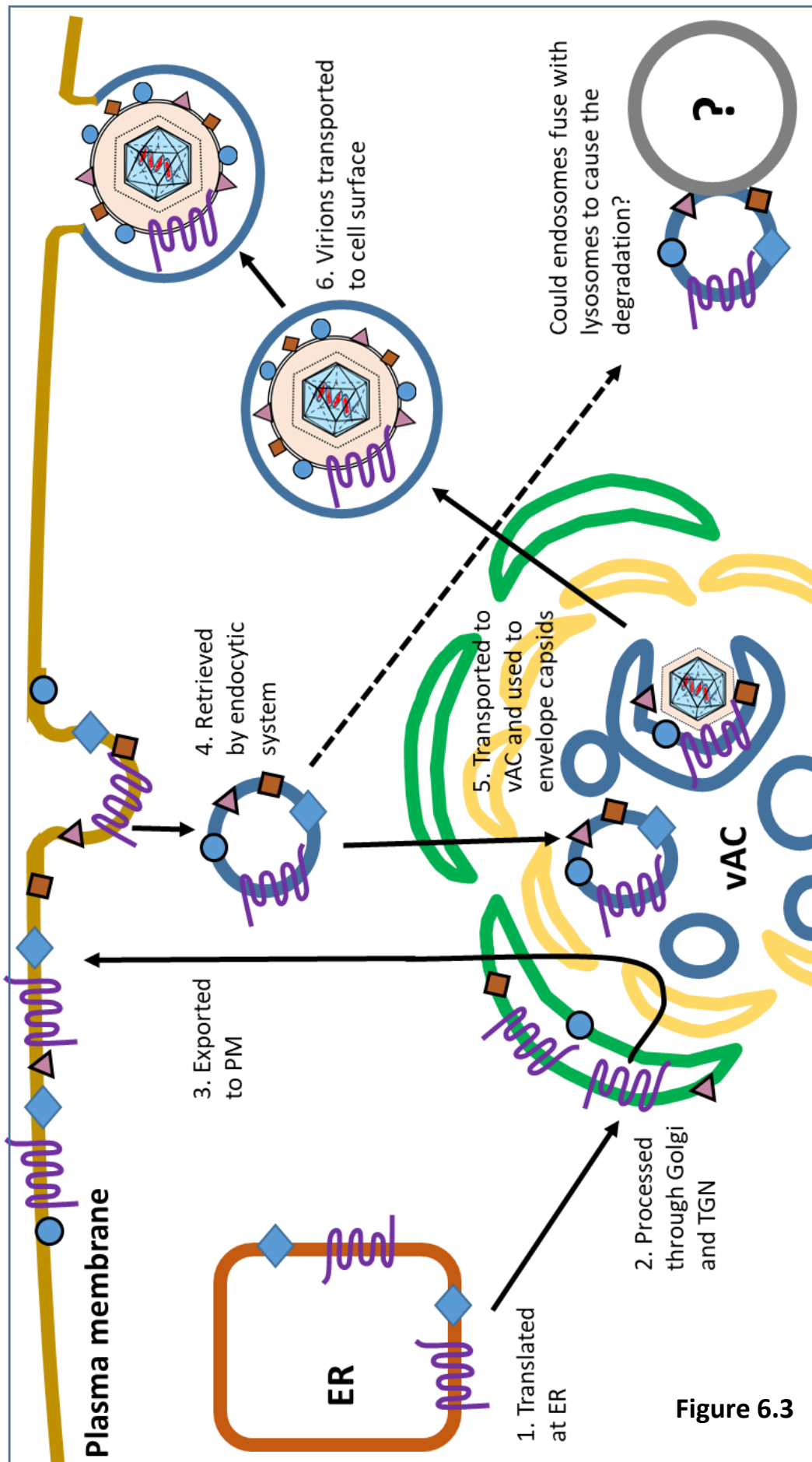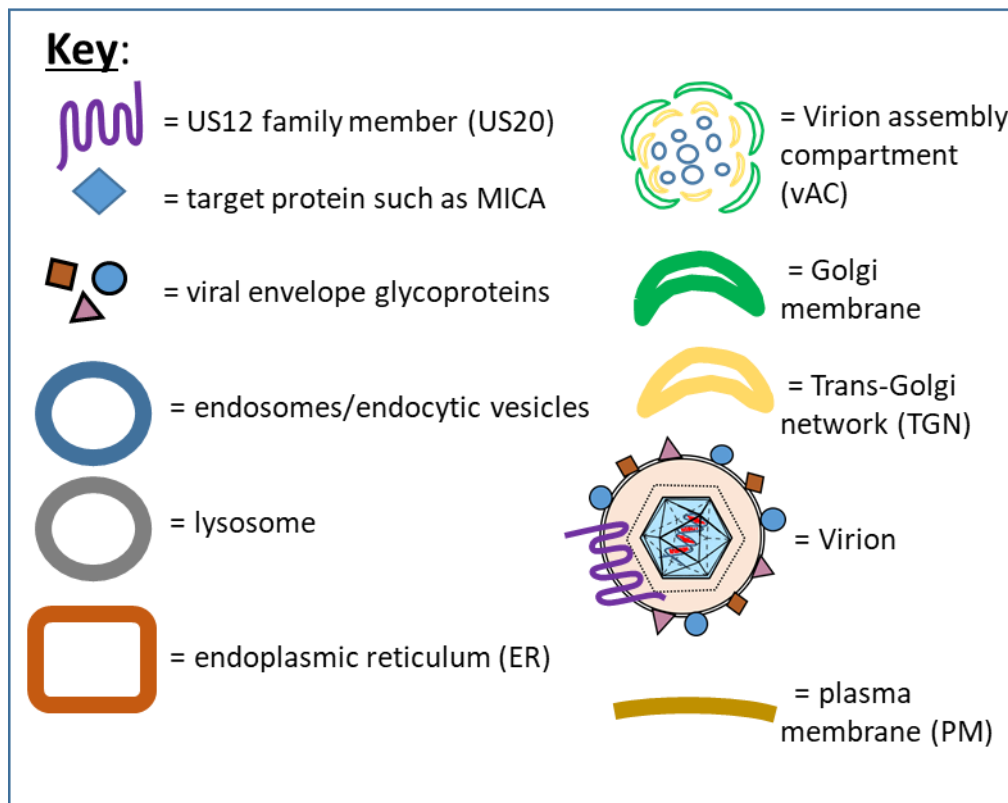
1. Translated at ER

Figure 6.3: Model for the hypothesis of US12 family members processing and transport through the cell. The US12 family members are likely translated at the ER, where some members are N-glycosylated. Then all or a subset of US12 family proteins are likely processed through the Golgi and TGN before being exported to the PM. The US12 family members may then be retrieved via the endocytic system, potentially along with their target proteins similarly to how virion glycoproteins are thought to be transported to the vAC. It is then likely that these endocytic vesicles (which accumulate in the centre of the vAC) can envelope the tegumented capsids and these enveloped virions are then egressed from the cell. It is possible that at some point during the transport cycle the endocytic vesicles can fuse with lysosomes in order to degrade their enclosed proteins which may account for the degradation of some US12 family members and their targets. In this diagram the US12 family are represented by US20 (purple 7TM protein) as US20 is known to localise to the ER and the vAC, has been identified on the PM by proteomics, and has been detected on HCMV virions. This however is likely not the case for all US12 family members, as not all US12 family members are found in the vAC or on the virion. The US12 family may also affect their target proteins at the ER.

are subsequently found in the centre of the vAC, however these endosomes can also fuse with lysosomes, which may be the mechanism through which the US12 family target proteins are degraded and would explain why US12 family members are often degraded themselves also (**Section 5.1**).

Evidence against this theory as well as the original hypothesis that the US12 family may transport their targets to lysosomes directly, includes the observation that US12 family members did not co-localise specifically with lysosomes, even in the presence of leupeptin. There could be a number of technical reason for this lack of co-localisation, as upon entering the lysosome the V5 tag may become shielded from antibodies as a consequence of modifications to the US12 family proteins, such as differential protein folding, obstructing interactions or the C terminus being trimmed. Although HCMV proteins have previously been shown to co-localise with lysosomes (Tirabassi and Ploegh, 2002), a positive control using a protein known to be targeted and degraded in the lysosome would have been beneficial in verifying whether immunofluorescence was accurate enough to conclusively detect specific co-localisation between proteins and lysosomes in this setting. However, as some co-localisation could occasionally be detected with some US12 family members (**Figure 5.5**), this suggests that the co-localisation would have been detectable. Therefore, the simplest explanation is that the US12 family of proteins are not being targeted to lysosomes, but act elsewhere. One possibility is that the US12 family could be directing their target proteins for lysosomal degradation without going there themselves and another possibility is that neither the US12 family or their targets are degraded lysosomally. Leupeptin is recognised to be a relatively specific lysosomal inhibitor due to its inhibition of serine and thiol proteases found in the lysosome, but more than 20% of its activity of serine and thiol proteases takes place outside of lysosomes (Seglen et al., 1979). Consequently, the leupeptin-sensitive proteolysis of US12 family members and/or their cellular targets could be non-lysosomal, and this would explain the lack of co-localisation of the US12 family with lysosomes. The expression of two US12 family target proteins MICA and B7-H6 however, could be rescued not only with leupeptin, but also by folimycin (concanamycin) which inhibits acidification of organelles, such as lysosomes (Fielding et al., 2014, Charpak-Amikam et al., 2017), increasing the likelihood of the degradation of at least some target proteins occurring in lysosomes. However, this would need further investigation and the use of additional specific protease inhibitors would need to be obtained to further specify the mechanism of action of the US12 family.

It would also be useful to investigate any effects of leupeptin on HCMV replication as although there is no published data pertaining to negative side effects of leupeptin on HCMV, other degradation inhibitors can have adverse effects. MG132 (a proteasome inhibitor) for example, is known to block viral DNA replication and assembly of HCMV (Kaspari et al., 2008). Therefore, an effect on HCMV replication or similar by Leupeptin cannot be ruled out, and experiments with Leupeptin should be analysed with caution. Observing the differences of HCMV vAC formation (using a vAC marker such as pp28) upon the addition of Leupeptin may provide an indication as to whether HCMV infection is affected by Leupeptin.

Other degradation pathways that cannot be rescued by leupeptin may also be utilised and the library of V5-tagged US12 family member HCMV's now permits for this to be investigated. As ubiquitin is one possible cause of the protein ladders seen on some US12 family members (**Section 6.1**), they may undergo ubiquitin-related degradation. Ubiquitination has previously been shown to target GPCRs for lysosomal or proteasomal degradation (Bonifacino and Traub, 2003), so could similarly be used to target the US12 family.

The degradation of US12 family protein targets was identified by proteomic analysis which revealed that many cellular proteins were downregulated on the PM and in the WCL, and many were demonstrated to be rescued by leupeptin (Fielding et al., 2017). A small subset however were downregulated on the PM, but not in the WCL, indicating that they may be retained, and possibly sequestered by the US12 family, rather than degraded. For instance, US13 downregulated MICA on the PM but not in the WCL and IL6ST was similarly regulated by US18. It would be interesting to investigate this retained subset of US12 family hits further by validating their cell surface downregulation by flow cytometry and using immunofluorescence to identify whether their cellular localisation matched that of the US12 family member responsible. If these proteins are sequestered, it could work on the same principle as HCMV gpUL141 which binds to CD155 (human PVR) in the endoplasmic reticulum and prevents its maturation and transport to the cell surface. It may also indicate that US12 family members work together to retain and then subsequently degrade their target proteins. For example, if one or more US12 family members targeted MICA for retention, and US13 was responsible for the degradation of US13, it would explain why MICA is downregulated from the PM but not the WCL in the HCMV US13 deletion mutant.

Many of the proteins that the US12 family have demonstrated to target are linked to the proposed or validated functions of those family members. For example, US12, US14, US18, US20 (and potentially US21) were demonstrated to have NK evasion functions, and US18 and US20 target many immune ligands and have been validated as targeting the NK ligands MICA and B7-H6 (Fielding et al., 2017, Fielding et al., 2014). The regulation of other NK ligands by the remaining family members has yet to be validated, however it appears that US12 regulates ULBP2 in the PM and WCL, and MICB in the WCL; and that US21 appears to regulate B7-H6 in the WCL (Fielding et al., 2017), however the potential NK evasion function of US21 may also be due to unforeseen effects on US20 by the deletion mutant. US14 has no known NK ligand targets (Fielding et al., 2017), so may be functioning in a different way to the other US12 family members that have NK evasion functions, and may possibly regulate a currently unknown or untested NK ligand. HCMV genes have predominantly undergone negative selection and are under strong evolutionary constraints, however US14 and US18 are subjected to higher levels of positive selection than would be expected from their diversity (Sijmons et al., 2015) (**Section 1.8.15**), suggesting that they were highly advantageous to the virus. This supports their proposed or proven roles in NK evasion, with many HCMV genes of high positive selection, predominantly functioning to modulate host immune and antiviral pathways (Sijmons et al., 2015). Although the whole family are dispensable for replication *in vitro*, the deletion of US13 from HCMV Towne does cause a moderate growth defect in fibroblasts (10^{-1}-10^{-2})(Dunn et al., 2003b). As there appears to be no unique immune ligands that US13 targets compared to the rest of the US12 family (Fielding et al., 2017), the growth defect may indicate an additional alternative function. The only noticeable difference within the target proteins of pUS13 is that MICB is more highly upregulated in the PM in the US13 deletion mutant than with the other US12 family member deletion mutants. Immunoblotting of immunoprecipitations (IP) of the V5-tagged family members would further identify which cellular proteins are binding to them, and could detect other potential target proteins. Not all target proteins may bind directly however so cellular targets that are indirectly regulated may remain unknown.

Targeting cellular proteins for degradation may help to reduce protein misfolding, leading to regulation of the ER stress response within the cell, and US17 can regulate the ER stress pathway through its regulation of ER stress response genes and chaperones, including BiP/GRP78, DDIT3/CHOP and CHAC1 levels (Gurczynski et al., 2014). BiP has many roles within the cell and is an ER chaperone that is involved in the folding and assembly of

proteins and can prevent the aggregation of unfolded proteins, modulating the regulation of the UPR response (Hegde et al., 2006). BiP can also bind calcium and is involved in maintaining calcium homeostasis in the ER (Buchkovich et al., 2008). TMBIMs can regulate the UPR response and apoptosis through the modulation of ER calcium homeostasis (Rojas-Rivera et al., 2012) so US17's regulation of BiP may also be linked with the modulation of calcium levels. BiP is upregulated 1.6 fold by HCMV AD169 ΔUS17, and each US12 family member deletion mutant in Merlin also appears to downregulate BiP (HSPA5) in proteomics (Fielding et al., 2017). Therefore, other members of the US12 family may also be able to modulate the ER stress response, and the finding that the majority of US12 family members are localised to the ER (**Section 5.3**) means that this interaction would be possible as BiP is also known to localise to the ER (Buchkovich et al., 2009). Additionally, BiP is later diverted to the vAC so the possibility of BiP chaperoning some of the US12 family members to the vAC should be investigated, although US12 family members may first need to translocate to the PM, especially those found on the virion. BiP is also required for virion assembly (Buchkovich et al., 2008) so its association with US17 may aid US17 in its role of modulating virion content (Gurczynski et al., 2014).

6.3 Localisation of the US12 family members within the cell

Immunofluorescence studies do have their limitations with some small differences seen across different experiments as indicated in **Table 5.2 (Section 5.6**). These small differences however are likely due to slight differences in the time of infection of each cell, with the 0 hpi time-point occurring after a 2 hour infection incubation period. The images presented in this thesis however are representative of at least 2 independent experiments and represent the most common localisations seen within each sample. Using percentages as a way of quantifying the different types of localisation concentrations within cells would have been useful to assess the representation in a quantifiable way. **Table 5.2** however indicates that the proteins have relatively consistent localisations between different experiments and the US12 family members usually only differ in their concentrations between their ER and vAC localisations within the same cell. If this was due to an impact of the timing of infection, it could potentially be partially alleviated in the future by synchronising the infection so that all cells are in the same growth phase when infected. Infecting the cells over a shorter time period than 2 hours (**Section 2.3.2**) may also aid in co-ordinating the time of infection, however the 72 hpi time-point may just be capturing the translocation of the US12 family (**Section 6.3.3**).

The localisation data on the US12 family presented here agrees with information available on members from the low passage strain TR and the highly passaged strain AD169. For both Merlin and TR, pUS16 localised to the outer portion of vAC (this thesis and Bronzini et al. (2012)) and gpUS20 associated with the ER (this thesis and Cavaletto et al. (2015)). In addition, Merlin pUS16-V5 and gpUS20-V5 could also occasionally be found in the ER and the vAC respectively. The trafficking of Merlin gpUS14-V5 to the ER and the vAC paralleled observations that AD169 gpUS14 was distributed in a uniform granular manner throughout the cytoplasm, concentrating in the vAC in some cells (Das and Pellett, 2007). HCMV Merlin pUS18-V5 could be detected in the cytoplasm at 72 hpi and was observed to marginally concentrate in the vAC and ER of some cells. Likewise, AD169 pUS18 exhibited a cytoplasmic localisation that became more concentrated in the vAC at later time-points up to 144 hpi (Das and Pellett, 2007). However, HCMV Merlin gpUS17-V5 exhibited an ER localisation which did not align with published data for AD169, in which gpUS17 was cleaved into 2 distinct domains with the N-terminus localised to the vAC and the C-terminus to the nuclei, the vAC and the cytoplasm (Das and Pellett, 2007, Das et al., 2006). The fact that the C-terminally tagged Merlin gpUS17-V5 was not detected in the nucleus suggests that Merlin gpUS17-V5 is unlikely to be processed or cleaved in the same way as AD169 gpUS17, and this is supported by the lack of nuclear localisation of Towne gpUS17-V5 also (Das and Pellett, 2007). The AD169 study also used BiP as the ER marker which has varying localisation with different BiP antibodies, and can be occasionally detected in the vAC (Buchkovich et al., 2009). HCMV infection may have also altered the distribution of the ER markers, with calnexin and BiP possibly not remaining in the same positions during HCMV infection. To formally address any differences, it would be useful to generate an antibody directed against the N-terminus of Merlin gpUS17, or attach an N-terminal tag and run a comparison by immunoblot, however it is not expected that the AD169 gpUS17 will represent the gpUS17 of clinical strains due to its highly passaged nature.

6.3.1 Endoplasmic reticulum localisation of US12 family members

ER localisation of US12 family members was demonstrated through their co-localisation with calnexin, a resident ER protein. The host cell machinery becomes rearranged during HCMV infection (**Section 1.4.1.6**), with the 'ER' becoming more condensed in HCMV infection than in uninfected cells, visualised as cytoplasmic structures towards the periphery of the cell (Cavaletto et al., 2015, Buchkovich et al., 2009, Alwine, 2012). The US12 family members' localisation to the 'ER' or ER-derived membrane would therefore

need verifying through the co-localisation of additional ER markers. Aside from calnexin (this thesis), gpUS20 has also been shown to co-localise with an additional ER marker, calreticulin (Cavaletto et al., 2015) which helps supports the location of gpUS20 as the ER.

All US12 family members, except for gpUS19-V5, showed at least partial co-localisation with the ER, and correspondingly were predicted to be translated at the ER, with transmembrane domains (TMDs) previously shown to act as localisation signals in the absence of signal motifs. Different TMD lengths have demonstrated to be the cause of the different localisations of syntaxins (Watson and Pessin, 2001), and altering the transmembrane domain of UL16 had previously shown to control its intracellular trafficking (Valés-Gómez and Reyburn, 2006). The ideal TMD for insertion into the ER is one of ~20 residues (Shao and Hegde, 2011) and the first TMD domains of the US12 family members are 22 residues long (except for US20 which is 21 residues long) so are all an appropriate length for localisation to the ER membrane. This hydrophobic TMD gets recognized by the signal recognition particle on the ribosome, at which point translation is halted until the ribosome moves to the ER (Shao and Hegde, 2011). This is known as co-translation and the vast majority of integral membrane proteins are assembled at the ER and inserted into the ER membrane as they are synthesized (Shao and Hegde, 2011). Being translated at the ER may also give them an advantage as it may reduce the competition of cellular mRNA translation. Each US12 family member (except for US18) also contained a consensus ER retention motif TRG_ER_diArg_1 or an N terminus XXRR-like ER motif; thus predicting their retention at the ER (**Section 4.1.2**). This may also explain why US18's co-localisation to the ER appears less frequent or more transient. Although the majority of US12 family members have at least a portion of their total protein 'retained' at the ER as predicted, another portion tends to associate with the vAC, although their mechanism of transfer to the vAC is unknown. They could achieve this translocation through the binding of other proteins bound for the vAC or may be chaperoned there. Proteomics suggests that some US12 family members are found on the plasma membrane also (Weekes et al., 2014, Fielding et al., 2017), so US12 family proteins may travel from the PM to the vAC via the endosomal transport system (**Figure 6.3**).

In accordance with their partial ER localisations, US12, US14, US15 and US16 were all identified in the ER/Golgi fraction of HCMV AD169 sub-organelle proteomics in at least 1 time-point across the course of infection (Jean Beltran et al., 2016). pUS12 was associated with the ER/Golgi fraction at 72 hpi, gpUS14 at 96 hpi, pUS15 at 96 and 120 hpi, and US16 at 120 hpi (Jean Beltran et al., 2016).

N-glycosylation findings also support the ER localisation demonstrated for gpUS14-V5, gpUS17-V5 and gpUS20-V5, with these protein species containing EndoH sensitive N-glycosylation, suggesting that they have yet to have passed through the ER to the Golgi. gpUS14-V5 was the only protein shown to be fully N-glycosylated and it had partial localisation to the ER. gpUS17-V5 was shown to be solely localised to the ER and showed only partial N-glycosylation, with a portion of gpUS17 protein remaining non-N-glycosylated. gpUS20-V5 also showed partial N-glycosylation but had dual localisations within the cell (the ER and vAC). This may mean that there is a chance that the N-glycosylated and un-glycosylated subsets may go to different locations within the cell, or that they have differing functions. Although N-glycosylation has only been demonstrated at 72 hpi, the band patterns of all 3 proteins remains the same over the course of infection, suggesting that this complete (gpUS14) or partial (gpUS17 and gpUS20) N-glycosylation is likely to occur across all time-points. gpUS20-V5 was the only member that was both N-glycosylated and also present in the virion, however it was unable to be determined whether the virion form of pUS20-V5 was similarly N-glycosylated. gpUS14 had previously been detected in the virion but this could not be verified in this thesis.

The sole ER localisation of gpUS17-V5 appears to preclude it from being packaged into virions as all members present in the virion have at least a partial or transient association with the vAC (**Section 5.3, 5.4 and 5.5**). gpUS17-V5 has been identified to function in virion maturation however, and other US12 family members involved in altering virion content or tropism also localise to the ER, and this may give them access to some of the virion components which are assembled there, or to manipulate virion proteins as they translocate through the ER. pUS16 for example is known to function by reducing the virion content of the pentamer on the virion (Luganini et al., 2017).

As mentioned there is some US12 family member homology to the TMBIM family (**Section 1.8.1.5**), and TMBIM proteins localize to membranes of various cellular organelles, including the ER (Lisak et al., 2015). At the ER, TMBIMs can modulate apoptotic and ER stress signalling by influencing cellular calcium levels (Lisak et al., 2015, Rojas-Rivera et al., 2012). US17 also appears to do this and therefore other members of the US12 family may function in a similar way and be involved in the regulation of ER stress pathways (as discussed in **Section 6.2**). pUS19 is the only member shown to have no co-localisation with the ER, and it is also the least similar to TMBIMs so is the least likely to be involved in these potential functions.

Confirming the topology and termini orientations of the US12 family experimentally would be helpful to fully determine their similarities and differences to TMBIMs (**Section 4.1.3, Figure 6.1**) and their structure would affect how they function. The predicted membrane topology for the US12 family (**Section 4.1.3**) is supportive of their N-glycosylation findings, as this orientation would allow for all N-glycosylation motifs of gpUS14 and gpUS17, and 2 of the 3 N-glycosylation motifs of gpUS20 to be present in the lumen of the ER and available for N-glycosylation. This orientation could be verified by performing selective permeabilization of the membranes during immunofluorescence to determine which side of the membrane the V5-tagged C-termini are on. The N-terminus could also be studied, however it would either need to be tagged separately or specific N-terminus antibodies would need to be used for each family member. Knowing the termini orientations would indicate at the orientation of all subsequent TMDs, although in order to specifically identify the orientation of the loop containing the N-glycosylation motif, you would need to use specific antibodies to that specific motif-containing segment would be required.

6.3.2 Localisation of US12 family members with the virion assembly compartment

All US12 family members, except for gpUS17-V5, show some association with the vAC, although they often do not show specific co-localisation to the UL99/pp28 vAC marker used. The vAC is made up of many host organelles, so further studies would need to be done to identify the specific vAC structure or membrane that each vAC-associated US12 family member is localised to, including co-staining with markers for the Golgi, trans-Golgi network (TGN) and early endosomes. If the US12 family co-localised with endosomes, this would further support the model that they may be trafficked from the PM to the vAC via the endosomal transport system (**Figure 6.3**). Having an association with the vAC means that these US12 family members could be in prime position to regulate the virion content, with pUS16-V5 known to regulate the amount of pentamer in the virion (Luganini et al., 2017). Other US12 family members that localise to the vAC may also have a function involved with the assembly and exit of virions, or may be present because they are added to the virion themselves. There does appear to be a clear trend between the vAC localisation of proteins and their presence in the virion, as demonstrated by the virion presence of HCMV vAC proteins such as pp28 (Sanchez et al., 2000b, Landini et al., 1987) and now with 7/9 of the vAC-localised members of the US12 family.

US12 family members that had only occasional or weaker localisation to the vAC at 72 hpi (pUS12, pUS13, pUS18 and pUS21), may instead become more concentrated at the vAC

later in infection, or their localisation in the vAC may be brief due to the rapid egress of the virions from the cell. This correlates to previous studies that indicated that AD169 pUS18 did not readily accumulate in the vAC until later time-points including 144 hpi (Das and Pellett, 2007). pUS16 was the only member detected in the virion preparations which was not abundant enough to be definitively designated as a virion protein, with its low level detection likely to reflect a background level of cellular contamination in virion preparations or passive low level incorporation into the virion due to its close association with the vAC, and its role in altering the virion content (Luganini et al., 2017). Other US12 family members found in the virion are present in relatively high abundance, which makes it much more likely that they have a function that requires being specifically packaged into the virion. For example, pUS18 and gpUS20 are both incorporated into the virion and both play a role in tropism (Cavaletto et al., 2015; Hai et al., 2006), and although their mechanism of action is currently unknown, is may require their presence in the virion. Equally, pUS18 and gpUS20 also have NK evasion functions and their presence in the virion may instead link to this role, and there is a chance that they could target their cellular NK ligands upon entry to the next newly infected cell. pUS12, gpUS14 and gpUS20 could be detected by QTV proteomics at 0 hpi (**Figure 4.2**) (Weekes et al., 2014) indicating that the presence of these US12 family members in the virion (this thesis and Murrell (2014)) was enough to lead to a detectable level of protein within the newly infected cell, which supports this theory. pUS12 and gpUS20 also appeared to be detected in the PM in cells infected with irradiated virus (Weekes et al., 2014), suggesting that their early delivery and localisation to the PM may be important for an early functional role. pUS12, gpUS14 and possibly pUS21 also have NK evasion functions and are also found on the virion, alongside other immune evasion proteins such as gpUL141 (this thesis and Murrell (2014)). It would be interesting to determine whether the levels in the virion would be high enough to utilise their immune evasion functions in the newly infected cell, and it would be interesting to see whether US18 or US20 for example could affect MICA levels solely from their presence on irradiated virus. The other members in the virion (pUS12-V5, pUS13-V5, pUS19-V5 and pUS21-V5) have unknown functions as of yet. They may also be involved in immune evasion, as they regulate an array of cellular immune ligands between them (Fielding et al., 2017), or their functions may be unrelated to immune evasion, and could be related to roles in tropism or similar. This is more likely for pUS13 and pUS19 which had the least number of cellular targets within the proteomics study, so are likely to be incorporated into

the virion for a role other than immune evasion. Some members may even be incorporated into the virion through their binding association with another virion protein.

As the US12 family are membrane proteins, it is expected that those present in virions will be found in the virion envelope. This hypothesis could be tested using electron microscopy (EM) with gold labelling of the V5-tagged viral proteins, as previously achieved for gpRL13-V5 (Stanton et al., 2010). Detergent fractionation of the virion could also be carried out by splitting the envelope into soluble (envelope) and insoluble (tegument and capsid) fractions to detect which fraction the US12 family proteins are present in, as previously achieved for US22 family members (Adair et al., 2002).

Both US18 and US20 are in the virion and both have two functions, NK evasion and tropism (Fielding et al., 2014, Cavaletto et al., 2015, Hai et al., 2006). Therefore other US12 family members may also have dual or multiple functions. US22 family members are tegument proteins found in the virion and they too are also implicated to function early in infection as well as having possible additional functions at later time points (Adair et al., 2002).

6.3.3 Translocalisation of US12 family members over time

Previous studies have shown differing US12 family localisations over time (Das and Pellett, 2007, Das et al., 2006), and this thesis identified US12 proteins in different concentrations or multiple locations at the same time-point. This suggested that either the 72 hpi time-point was capturing a translocation period between the vAC and ER, or that dual localisation may be seen across the time-course of infection. Therefore, it would be beneficial for future studies to observe the proteins at multiple time points throughout the course of infection and determine if and how their localisations change. If the 72 hpi time-point is capturing a change in localisation over time, it would be beneficial to know which direction they are translocating and whether family members remain in the same localisations as each other across all time-points. The localisation of US12 family members (**Section 5.3 and 5.4**) are likely to link to their functions (**Section 1.8.15**) and their translocation likely also reflects this. If multiple localisations are common across all time-points of infection, this may indicate that US12 family members could have multiple functions, or may need to go to multiple locations to perform a single function. This would correlate with the broad range of target proteins that members have, especially US20 which targets 54 cellular proteins (Fielding et al., 2017), and links to the fact that multiple members have already been shown to have more than 1 function, including US17, US18

and US20 (Gurczynski et al., 2014, Cavaletto et al., 2015, Hai et al., 2006, Fielding et al., 2014, Fielding et al., 2017). As the ER stress pathway (and BiP) have also been shown to have a role in virion maturation, assembly complex formation and function (Buchkovich et al., 2008) this potentially links both the ER and vAC localisations of the US12 family, and their roles could potentially modulate pathways related to protein folding or trafficking between the ER and the vAC. Some US12 family proteins may solely travel to the vAC to be packaged into virions (**Section 5.5**).

The majority of US12 family members are found in the ER and/or the vAC, with multiple members targeting the same cellular proteins, with some US12 family members having more dispersive vAC associations and others having more distinctive vAC patterns. There are no obvious links between the specific localisations of each US12 family protein and the proteins that they target, with US18 and US20 targeting many of the same proteins, without appearing to be in large concentrations at the same localisations as each other. This could potentially mean that they have slightly different ways of directing the same protein for degradation, or are able to direct their targets for degradation in the same way from their respective locations. It does however appear that they would overlap in small quantities in the ER and the vAC, and may co-localise more specifically at different time-points of infection. Although pUS19-V5 showed association with the vAC, it was unique in that it ringed around the vAC periphery in a distinctive pattern. It is also the member that targets the least number of proteins in common with other family members, and is one of the few members not affected by the addition of leupeptin. Combined with US19 having the least homology to GPCRs and the rest of the US12 family (Lesniewski et al., 2006), it is likely that US19 has a different function to the other family members.

It may be possible that only the location(s) in which the US12 family proteins are most abundant can be clearly visualised by immunofluorescence. This may explain why some members are found at a single site within one cell, and have multiple localisations in a neighbouring cell. This may also account for why US12 family members cannot be visualised at the plasma membrane, despite being detected in the PM by proteomics (Weekes et al., 2014, Fielding et al., 2017), else their association with the PM may be too transient, with the proteins being rapidly translocated to the vAC. Interestingly, the fact that the C-termini are predicted to be non-cytoplasmic, means that if the US12 family members are on the surface, their C-termini would be extracellular (**Figure 6.1**) and the opportunity then exists to use the V5-tag in flow cytometry to directly measure whether any US12 family member can be detected on the cell surface. Their presence on the plasma

membrane would be logical for members found on the virion as virion envelope proteins go to the cell surface to form the virion envelope before going back to the vAC to envelope the nucleocapsids (**Section 1.4.1.7**).

In considering the impressive role that the US12 family plays in modulating surface expression of immune ligands/receptors, it seemed reasonable to expect that US12 family members would be found associated with their target proteins, as has been observed for other HCMV immune evasion functions such as gpUL141 (Prod'homme et al., 2010, Smith et al., 2013b, Tomasec et al., 2005). For many of the US12 family targets such as MICA, B7-H6 and ALCAM, it proved difficult to source antibodies that are fully compatible with the immunofluorescence technique (**Section 5.6**). The fact that many of these targets are also targeted by proteolysis by the actions of the US12 family does not help such analyses, although the addition of inhibitors such as leupeptin could be used to rescue the targets. An antibody to MPZL1 did however give a defined localisation, and appeared within the vicinity of the vAC (**Section 5.6**) and it could be subsequently tested for its co-localisation with gpUS14, pUS18, and gpUS20 which have all been implicated in targeting MPZL1 (Fielding et al., 2017). If these US12 family members do not chaperone their target proteins for degradation however, then their association with them may be transient and may still be hard to detect through immunofluorescence. Nevertheless, it is now possible to address this and other interesting degradation and co-localisation questions using the constructed HCMV library of V5-tagged US12 family gene members.

6.4 The possibility of the US12 family working in complexes

pUS18 and gpUS20 have already been demonstrated to work in concert, with US18 and US20 targeting MICA for degradation at a much higher rate than either member individually, with US18 only marginally affecting MICA levels alone (Fielding et al., 2014). They were similarly shown to act in concert to suppress B7-H6 to reduce NK activation, and both target multiple other cellular proteins including ULBP2, IL6ST, KIT, KITLG, JAM3, ACVR1, ACVRL1, IFNGR1, MPZL1, CXADR, ALCAM, SDC4, CD99 and SDC1. A subset of these cellular targets were additionally regulated by US14 and/or US16 (Fielding et al., 2017). In fact, across all plasma membrane proteins targeted by the US12 family, 29% were regulated >3 fold by 2 or more family members, and 6% targeted by 3 or more family members (Fielding et al., 2017). US12 family members may therefore work together in a complex. The deletion of certain family members also affects the expression levels of other

members, with US14 expression increased during infection with HCMV ΔUS15, and US15 expression increased during infection with HCMV ΔUS20 (Fielding et al., 2017). This suggested that US12 family members may compensate for each other, further implicating a cross-over of functions.

Multiple US12 family members are also found in the same general localisations as each other, with 9/10 members co-localising with the ER marker calnexin (**Section 5.3**), implying that all of these members may be in close enough proximity to form a complex. 9/10 of the US12 family are also associated with the vAC, with the majority of members generally showing dispersive vAC association (**Section 5.4, Figure 6.4**). This was true for pUS12-V5, pUS13-V5, gpUS14-V5, pUS15-V5, gpUS20-V5 and pUS21-V5, with pUS18-V5 showing occasional concentration in the vAC also, so in theory these members would be physically close enough in the vAC to be able to complex. Although occasionally dispersive in the vAC, pUS16-V5 generally localised to the outer portion of the vAC only, and pUS19-V5 was observed in a distinctive pattern around the periphery of the vAC (**Section 5.4, Figure 6.4**). Due to this unique localisation of pUS19-V5, this member was the least likely to form a complex with any other US12 family members. It would be beneficial to further specify the exact vAC localisations of each US12 family member with other vAC organelle markers.

A large subset of members were additionally regulated by leupeptin, so appear to be targeted for degradation in the same way. Although this does not prove that they are working together, it does suggest that they are functioning in a similar way. To more rigorously test this US12 family complex hypothesis, multiple members could be tested for co-localisation with each other. We have recently acquired antibodies specific to pUS18 and gpUS20 so after optimisation, these antibodies could be used to test the co-localisation of all V5-tagged US12 family members with pUS18 and gpUS20. Multiple US12 family members could also be tagged with different tags (e.g. the V5 tag, Strep tag and 6xHis tag) within the same virus to observe co-localisations. Confocal or super resolution microscopy could also be exploited to enhance the resolution which may further inform on potential co-operation between family members.

Stable isotope labelling with amino acids in cell culture (SILAC) immunoprecipitations (IPs) could also be undertaken to identify each V5-tagged family member's binding partners by mass spectrometry. During SILAC IP, samples are differentially labelled with heavy, medium or light isotopes which allows the V5-tagged sample to be directly compared to the non-tagged Merlin control in order to identify any differences in protein enrichment.

Some members may complex together in membranes

ER

Cytoplasm

Outer vAC (Golgi and TGN)

Lysosomes

vAC

Figure 6.4

Nucleus

Key:

= US12 = US17 = vAC (as depicted in **Figure 1.7**)

= US13 = US18

= US14 = US19 = ER membrane

= US15 = US20

= US16 = US21 = lysosomes

→ All US12 family members are 7TM so will be present through a membrane however these membranes have not been depicted (except for the ER)
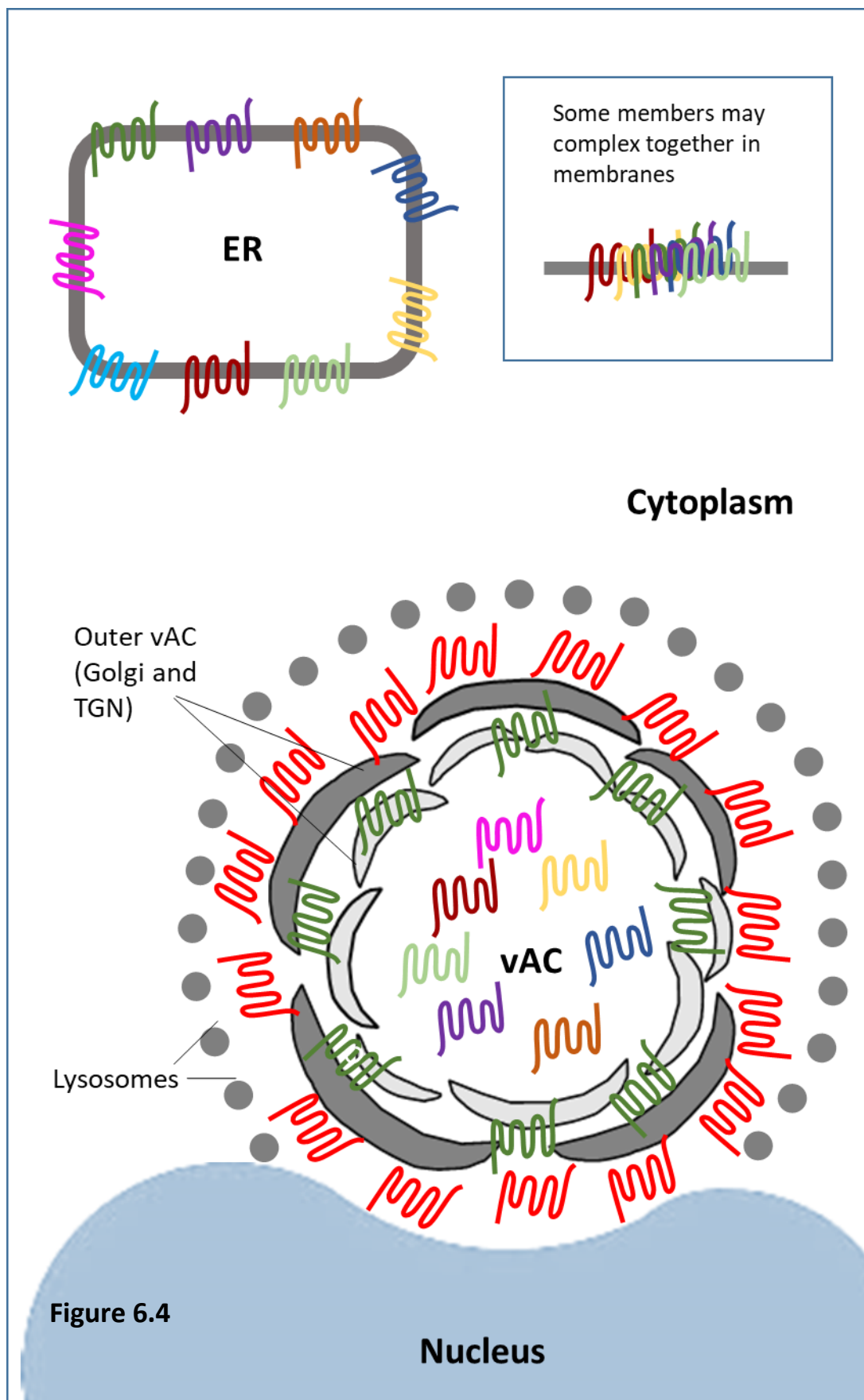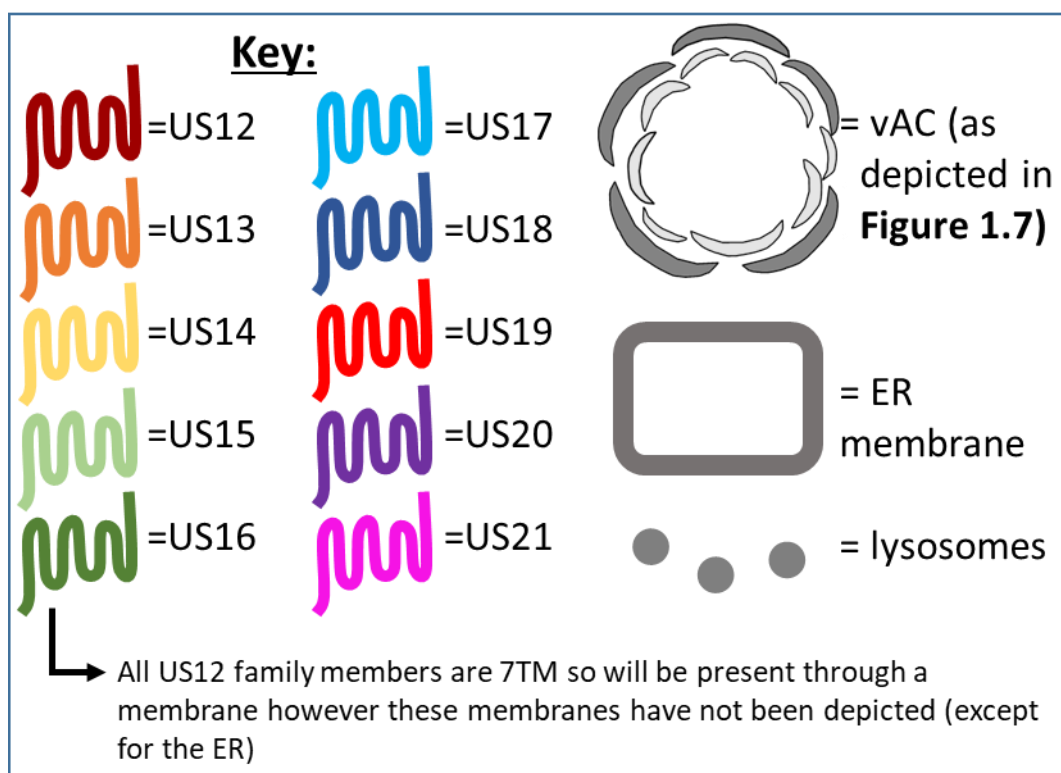
Figure 6.4: Depiction of the localisations of the US12 family within the cell at 72 hpi. Observed localisations of US12 family members within the cell in relation to the ER, the vAC (as crudely depicted by concentric circles of the Golgi (dark grey) and TGN (light grey), as portrayed in more detail in **Figure 1.7**), and the lysosomes that were demonstrated to ring around the vAC in HCMV infected cells. Localisations of each US12 family member were identified across multiple immunofluorescence experiments at 72 hpi (**Table 5.2**) and show the multiple localisations for some US12 family members, however their concentration and frequency of localisation to each location has not been depicted. US12 family members are depicted as 7 transmembrane protein patterns (**Figure 4.1**) to represent the crossing of each protein through a membrane, however specific membranes within the vAC have not been identified and therefore could not be depicted. pUS12-V5, pUS13-V5, gpUS14-V5, pUS15-V5, pUS18-V5, gpUS20-V5 and pUS21-V5 have dispersive vAC localisations, whereas pUS16-V5 tends to localise more to the outer portion of the vAC, and pUS19-V5 has a unique pattern around the periphery of the vAC. All members except pUS19-V5 co-localise to the ER however many not necessarily be present within the same membrane as depicted. Some US12 family members may also complex together (**Section 6.4**). All US12 family members are 7TM proteins so they will be present through a membrane, although these are currently unidentified (except for the ER).

Preliminary SILAC IP studies on pUS18-V5 and gpUS20-V5 indicated that multiple US12 family members are 'pulled down' in the IP, through their direct or indirect binding with pUS18-V5 or gpUS20-V5. This study identified that gpUS14, pUS15, pUS16, pUS18 and gpUS20 were typically detected in both pUS18-V5 and gpUS20-V5 IPs (Dr C. Fielding). This suggests that these members are likely to co-operate and possibly complex together, and undertaking SILAC IPs on all remaining V5-tagged US12 family members will help to further solidify this theory. It may also reveal new binding partners within the US12 family. SILAC IP data suggests that gpUS14, pUS15, pUS16, pUS18 and gpUS20 may form a complex together or can bind to each other, and this is largely supported by immunofluorescence data. All of these members have at least partial association with the ER and the vAC, and pUS16-V5 is the only member that appears to have an altered localisation within the vAC, although it is possible that it is also dispersive across the vAC in a lower abundance that was not easily detected by immunofluorescence. It is interesting to speculate that the proposed complex may form in one or both of these localisations and it may be possible that the members of the complex could change depending on the conditions within the cell, or have different associated members at different locations within the cell. Other complexes have also been known to alter their constituents, including the entry complexes, which can vary between gH/gL/gO and gH/gL/pUL128-131 (Li et al., 2015, Luganini et al., 2017, Zhou et al., 2013).

These potential complex members all belong to expression class Tp2, Tp3 or Tp4, all with relatively high expression levels at 48 and 72 hpi, which would aid their ability and likelihood of being able to complex together. It is unlikely that pUS19 forms part of this complex, due to its different localisation, and its lack of crossover with the rest of the US12 family's cellular targets (this thesis and Fielding et al. (2017). pUS19-V5 was also not rescued by the addition of leupeptin (**Section 5.1**) and has the least similarity to TMBIM family members (Lesniewski et al., 2006). Although co-operation between HCMV proteins is not novel, with HCMV US2 and UL141 known to work together (Hsu et al., 2015), multiple members of a family or a whole HCMV gene family working together in a complex has not been previously described or inferred. Working in complexes would further explain why the US12 family members are so difficult to study in isolation. It also gives limitations to the results recovered from single deletion mutant assays as the deletion of one member may not detect all of its target proteins if there is redundancy in function. This was first recognised with US18 which had limited effects on MICA when deleted individually (Fielding et al., 2014). The block deletion mutant therefore gives extra insights into proteins

that may be targeted by multiple members at once by showing a larger increase of a target protein than what can be explained by single genes together (Fielding et al., 2014, Fielding et al., 2017). However, without double and triple knockout mutants of multiple family members, the full extent of the families targeting of cellular proteins and which members of the family operate together may not be fully realised.

6.5 Conclusions

All US12 family members were successfully tagged within the context of HCMV and all were detected by immunoblot which was an advantage over previous methods for detecting the US12 family. A total of nine US12 family members were shown to co-localise to the ER (or ER-derived membranes) and nine family members showed association with the vAC, which corresponded to 7 of those family members being detected in the virion. This data corresponds to the known functions of the family including their involvement with tropism and virion content modulation, along with their NK evasion functions. This thesis has also shed light onto the fact that the US12 family members are degraded similarly to their targets, and that this degradation is unlikely to be solely lysosomal. The importance of studying HCMV proteins within the context of infection, especially 7TM proteins, has also been re-affirmed.

It has previously been suggested that the US12 family should be split into sub-families (Lesniewski et al., 2006; Rigoutsos et al., 2003) with US17, US18, US19 and US21 suggested to be separated from the family definition altogether. However, this data provides evidence to the contrary, given that the family all have the same localisations, and that multiple members target the same cellular proteins, especially US18 and US20, so it is highly unlikely that they are not related. It is also possible that some US12 family members may work together in complexes and this thesis has provided further supporting evidence of this hypothesis.

Gene families as a whole are generally not studied together, so it has been beneficial to observe all members at once in terms of their expression and localisations within the same system to get a better idea of how they work. Studying the US12 family members in parallel has therefore furthered previous studies that were done across a range of strains and cell types on just a few family members. Some members such as US12, US13, US15, US19 and US21 had no previous experimental data published on them in terms of their expression and/or localisations. This study has revealed both similarities and differences between

US12 family members and resulted in further characterisation of this gene family. As a whole this study has identified many novel traits of the US12 family, including 2 novel N-glycosylated proteins, novel ER and vAC localisations within the cell, previously unknown expression patterns over time, and the novel characteristic that over half of the family are targeted for degradation.

7 Appendix

Appendix Table 7.1: List of primers and oligos used during the recombineering, PCR and sequencing stages of cloning

Gene	Primer type	Sequence
US12	SacB/Rpsl cassette insertion	**For:** CGTCGGGAGAACACGGTGTTTTAGGGTGCGGGGGACAAAGGACAGTA CGACAGATTAGGTGATAGAAACGTTTTTTTTTACCTGTGACGGAAGATCACTT **Rev:** TAGTACCCCTGACGGCCCCCATCTGGTATCCAAACTACGCCGGGGCCCT AGGCCGCACGGCACACTGGCTTTTTCATAAACTGAGGTTCTTATGGCTCTTG
	Sequencing primers	**For:** CCCTGTCTAGACTCAAAAG **Rev:** ATCGTCCCCCTTTCTCTATA **Int Rev:** GCTAAGTTTCATGCTTCC
	V5 tag oligos	**For:** GCGGGGGACAAAGGACAGTACGACAGATTAGGTGATAGAAACGTTTTTTT ttacgtagaatcaagacctaggagcgggttagggattggcttaccagcgct **Rev:** CAAACTACGCCGGGGCCCTAGGCCGCACGGCACACTGGCTTTTTCATAAA Agcgctggtaagccaatccctaacccgctcctaggtcttgattctacgtaa
US13	SacB/Rpsl cassette insertion	**For:** CGGGTGCTCGACGAACAGTCGTCGGGGCTTCAGGTACCCGGCAAGTTTTA TAGAGAAAGGGGGACGATGGGTGGTGGCTACCTGTGACGGAAGATCACTT **Rev:** GTCCCAACCTCTGGCGCCTGCCCTGGACGACCGTCTTTGCCGCCTTCAG ATCCTCGTATTGCGAAGGTGGCGGTGGCTCGCTGAGGTTCTTATGGCTCTTG
	Sequencing primers	**For:** GCCGAGTGGCTCGCC **Rev:** CTGGGCACCTATCATCATTA **Int Rev:** ATGGTTGGGGACAGTTTT
	V5 tag oligos	**For:** TTCAGGTACCCGGCAAGTTTTATAGAGAAAGGGGGACGATGGGTGGTGG ttacgtagaatcaagacctaggagcgggttagggattggcttaccagcgct **Rev:** CGTCTTTGCCGCCTTCAGATCCTCGTATTGCGAAGGTGGCGGTGGCTCG agcgctggtaagccAatccctaacccgctcctaggtcttgattctacgtaa
US14	SacB/Rpsl cassette insertion	**For:** GGGTCCATGAGGCGGGTGATGCGCCCGAGTGAACGGGTGAGCGTCTC GGTGGAGTCTTCTTATAAACCAGCGGGTCTCACCTGTGACGGAAGATCACTT **Rev:** ATCATCATTACCGGGGTTGACAACGGCACGCTCAGCGTCATCCTCAACAG CACCACCGCGACGTTCCAGAGCAGGGTTGCTCTGAGGTTCTTATGGCTCTTG
	Sequencing primers	**For:** GGAGGGAAGCCCATTGC **Rev:** TCATTACCTGTCTAGCCG **Int Rev:** ACCTGGTTGCATAAGACT
	V5 tag oligos	**For:** AGTGAACGGGTGAGCGTCTCGGTGGAGTCTTCTTATAAACCAGCGGGTC ttacgtagaatcaagacctaggagcgggttagggattggcttaccag **Rev:** CAGCGTCATCCTCAACAGCACCACCGCGACGTTCCAGAGCAGGGTTGCT Agcgctggtaagccaatccctaacccgctcctaggtcttgattctacg
US15	SacB/Rpsl cassette insertion	**For:** CGCGGTTTCCCGCTGCGTGGAAACTGTCTCCATGTCGGGACCGCAGCG CCCGGCGGCGTATCCGCAAGGTCTCGAAGCTACCTGTGACGGAAGATCACTT **Rev:** TGGTTTTCACTCTGCTGATGGTGCTGAGAATCATGACCCTGCGCACCTTT

		TTGCAAACCTACTTTTCCTCTGACAAGCTGCTGAGGTTCTTATGGCTCTTG
	Sequencing primers	**For:** CGGACGCGGCTTCC **Rev:** GTCGCTACAGCTCTTTATTA **Int Rev:** CACCTTACTGGCCTTTCT
	V5 tag oligos	**For:** TCCATGTCGGGACCGCAGCGCCCGGCGGCGTATCCGCAAGGTCTCGAAG ttacgtagaatcaagacctaggagcgggttagggattggcttaccagcgct **Rev:** CATGACCCTGCGCACCTTTTTGCAAACCTACTTTTCCTCTGACAAGCTG Agcgctggtaagccaatccctaacccgctcctaggtcttgattctacgtaa
US16	SacB/Rpsl cassette insertion	**For:** CCCTTTTTCTCTTCTCATGGTGCGCTGCGTTCTCTGGAAACGGCTGCTCT GTCCGAAAACCAGTTCCGAACGAAAATCTACCTGTGACGGAAGATCACTT **Rev:** GGGCGAGAGGGTGGACAACGGCGTTGACGACGAAGCATGGGACAGGT CGTTCGGCGTTAACGTCATCGCGTCGGACGACGCTGAGGTTCTTATGGCTCT TG
	Sequencing primers	**For:** GGGGCACGTAGATGACCG **Rev:** CTCATTAGACAAACTCATCG **Int Rev:** TCGTGGTCTTTCTGGCTA
	V5 tag oligos	**For:** GTTCTCTGGAAACGGCTGCTCTGTCCGAAAACCAGTTCCGAACGAAAAT ttacgtagaatcaagacctaggagcgggttagggattggcttaccagcgct **Rev:** CGACCTGTCCCATGCTTCGTCGTCAACGCCGTTGTCCACCCTCTCGCCC agcgctggtaagccaatccctaacccgctcctaggtcttgattctacgtaa
US17	SacB/Rpsl cassette insertion	**For:** GGCGGCCCGCGGTTCTAACAGGCTTGATTGGTGGAGACGGCCGGCGCGG CGGGTGGGGGAAACGACGAGTTTTTCCGTTACCTGTGACGGAAGATCACTT **Rev:** CGCCATGGTTCGCGTGAGGTTTCTCTGTACCTCCCGCAAAAGGTCACAGC CCGAAATGGAGGCCGCGTTGGTGGCCCCGGCTGAGGTTCTTATGGCTCTTG
	Sequencing primers	**For:** GTCTAAGACGCGAGATCCG **Rev:** CCCAGTAGACAGACAGAACA **Int Rev:** GGGCCTGCTCACCATTTA
	V5 tag oligos	**For:** TGGTGGAGACGGCCGGCGCGGCGGGTGGGGGAAACGACGAGTTTTTCCG ttacgtagaatcaagacctaggagcgggttagggattggcttaccagcgct **Rev:** CTGTGACCTTTTGCGGGAGGTACAGAGAAACCTCACGCGAACCATGGCG Agcgctggtaagccaatccctaacccgctcctaggtcttgattctacgtaa
US18* (Fielding et al., 2014)	SacB/Rpsl cassette insertion	**For:** GGGAGGTTCATCGTCTGTCTCTAGAGGGAAGGTGGGGAACGTCTAAGCG AGCGGGAGCGTGTCATCTCCCCCATCTTTCCTGTGACGGAAGATCACTTCG **Rev:** CGGCCACGTCTGGGTGCAGCAGTACGCCGAGAAACACGGCGGACGCATC GACGGCGTGAGTCTCCTCAGCTTGTTGTAACTGAGGTTCTTATGGCTCTTG
	Sequencing primers	**For:** AGAGTGTAATATAATCACCG **Rev:** CTCTATGTCGAAAATGTGGC
	V5 tag oligos	**For:** AGGTGGGGAACGTCTAAGCGAGCGGGAGCGTGTCATCTCCCCCATCTTT ttacgtagaatcaagacctaggagcgggttagggattggcttaccagcgct **Rev:** CGAGAAACACGGCGGACGCATCGACGGCGTGAGTCTCCTCAGCTTGTTG Agcgctggtaagccaatccctaacccgctcctaggtcttgattctacgtaa

US19*	SacB/Rpsl cassette insertion	**For:** CAGCACCCGGTTACCGCGGATTTGATTGACGTCACGAGTGTGGTCAAACCG TGGCGGCACCCTGTATCCGACCCGTCG<u>CCTGTGACGGAAGATCACTTCG</u> **Rev:** GCTACGCCTCTATGTCGAAAATGTGGCTTTATTCATCGGCATGTACCATCTT CTGAGGCTCTGGTTGTGGAGCCCATGA<u>CTGAGGTTCTTATGGCTCTTG</u>
	Sequencing primers	**For:** GGAGCGGCACGATGGTGACC **Rev:** TCTGCCCACCTAACCAATGC
	V5 tag oligos	**For:** CGTCACGAGTGTGGTCAAACCGTGGCGGCACCCTGTATCCGACCCGTCG ttacgtagaatcaagacctaggagcgggttagggattggcttaccagcgct **Rev:** TTTATTCATCGGCATGTACCATCTTCTGAGGCTCTGGTTGTGGAGCCCA Agcgctggtaagccaatccctaacccgctcctaggtcttgattctacgtaa
US20* (Fielding et al., 2014)	SacB/Rpsl cassette insertion	**For:** ACGGTCCATTCTAGCGGGACGACATGAAGCATGGCGACAAGCGCGGCTG CTGTGAAAACGGGCGCGGTTTTATAGGCA<u>CCTGTGACGGAAGATCACTTCG</u> **Rev:** CCGTTGGATTAGTCTTTCGGACGGCGCGCCTTTGGACAACGGGACTTT GACAGCCGCCAGTACGACGGGGAAGTCCTAA<u>CTGAGGTTCTTATGGCTCTTG</u>
	Sequencing primers	**For:** TAGCTCGGCCACCGGTGGCG **Rev:** TCCGTGCTCTACTTCATGCC
	V5 tag oligos	**For**: CATGGCGACAAGCGCGGCTGCTGTGAAAACGGGCGCGGTTTTATAGGCA ttacgtagaatcaagacctaggagcgggttagggattggcttaccagcgct **Rev:** GCCTTTGGACAACGGGACTTTGACAGCCGCCAGTACGACGGGGAAGTCC Agcgctggtaagccaatccctaacccgctcctaggtcttgattctacgtaa
US21*	SacB/Rpsl cassette insertion	**For:** TGCGGCGCACCTACCCTTCTCTTATACACAAGCGAGCGAGTGGGGCACG GTGACGTGGTCACGCCGCGGACACGTCGA<u>CCTGTGACGGAAGATCACTTCG</u> **Rev:** CAGCGCCCACACTGCTCAGACGACGGTCGCTGCGACGGTCGCTGCCACA GCAGCGGCGTCGCCCCAGTTCGTCTCCTAA<u>CTGAGGTTCTTATGGCTCTTG</u>
	Sequencing primers	**For:** GCTGAAAGATGAAGATGGCG **Rev:** ACCCGACCAGATGGGAGACG **Int Rev:** GTCAGGCTTCCACTTTAG
	V5 tag oligos	**For:** AAGCGAGCGAGTGGGGCACGGTGACGTGGTCACGCCGCGGACACGTCGA ttacgtagaatcaagacctaggagcgggttagggattggcttaccagcgct **Rev:** CGCTGCGACGGTCGCTGCCACAGCAGCGGCGTCGCCCCAGTTCGTCTCC Agcgctggtaagccaatccctaacccgctcctaggtcttgattctacgtaa

For= Forward primer. **Rev**= Reverse primer. **Int Rev**= Internal reverse primer. Lowercase letters denote the V5 tag portion of the primer and underlined letters denote the SacB/Rpsl cassette homology

*US18-V5 and US20-V5 had been completed, and mid-stage clones of US19-Rpsl and US21-Rpsl had been prepared (Dr Ceri Fielding).

Appendix Table 7.2: The guanine-cytosine (GC) content of US12 family members

Gene	NCBI† Gene ID	NCBI† Aliases	GC content (%)*
US12	3077562	HHV5wtgp143	55.79
US13	3077576	HHV5wtgp144	57.38
US14	3077456	HHV5wtgp145	59.49
US15	3077565	HHV5wtgp146	57.67
US16	3077558	HHV5wtgp147	59.78
US17	3077567	HHV5wtgp14	58.39
US18	3077472	HHV5wtgp14	57.94
US19	3077522	HHV5wtgp150	62.79
US20	3077561	HHV5wtgp151	54.64
US21	3077437	HHV5wtgp152	55.87

† The National Center for Biotechnology Information (NCBI)

* GC content redicted by a DNA/RNA GC content calculator website

(http://www.endmemo.com/bio/gc.php) using DNA sequences taken from HCMV Merlin

(Human herpesvirus 5 -NC_006273.2) on NCBI. Percentages rounded to the nearesr 0.01%

Appendix Table 7.3: Clone numbers, designated virus codes for each once fully sequenced

Gene tagged	Clone identifier	Virus number
US12-V5	Clone 1-2	RCMV2314
US13-V5	Clone 4C	RCMV2172
US14-V5	Clone 4C	RCMV2174
US15-V5	Clone 2F	RCMV2190
US16-V5	Clone 16	RCMV2329
US17-V5	Clone 7	RCMV2330
US18-V5	Clone 2	RCMV1692
US19-V5	Clone neo1-1	RCMV2158
US20-V5	Clone 3	RCMV1691
US21-V5	Clone 6A	RCMV2192
None	Parental virus	RCMV1111

References

ADAIR, R., DOUGLAS, E. R., MACLEAN, J. B., GRAHAM, S. Y., AITKEN, J. D., JAMIESON, F. E. & DARGAN, D. J. 2002. The products of human cytomegalovirus genes UL23, UL24, UL43 and US22 are tegument components. *J Gen Virol,* 83**,** 1315-24.

ADLAND, E., KLENERMAN, P., GOULDER, P. & MATTHEWS, P. C. 2015. Ongoing burden of disease and mortality from HIV/CMV coinfection in Africa in the antiretroviral therapy era. *Frontiers in Microbiology,* 6**,** 1016.

ADLER, S. P., MANGANELLO, A. M., LEE, R., MCVOY, M. A., NIXON, D. E., PLOTKIN, S., MOCARSKI, E., COX, J. H., FAST, P. E., NESTERENKO, P. A., MURRAY, S. E., HILL, A. B. & KEMBLE, G. 2016. A Phase 1 Study of 4 Live, Recombinant Human Cytomegalovirus Towne/Toledo Chimera Vaccines in Cytomegalovirus-Seronegative Men. *J Infect Dis,* 214**,** 1341-1348.

ADLER, S. P., PLOTKIN, S. A., GONCZOL, E., CADOZ, M., MERIC, C., WANG, J. B., DELLAMONICA, P., BEST, A. M., ZAHRADNIK, J., PINCUS, S., BERENCSI, K., COX, W. I. & GYULAI, Z. 1999. A canarypox vector expressing cytomegalovirus (CMV) glycoprotein B primes for antibody responses to a live attenuated CMV vaccine (Towne). *J Infect Dis,* 180**,** 843-6.

ADLER, S. P., STARR, S. E., PLOTKIN, S. A., HEMPFLING, S. H., BUIS, J., MANNING, M. L. & BEST, A. M. 1995. Immunity induced by primary human cytomegalovirus infection protects against secondary infection among women of childbearing age. *J Infect Dis,* 171**,** 26-32.

AHLQVIST, J. & MOCARSKI, E. 2011. Cytomegalovirus UL103 Controls Virion and Dense Body Egress. *Journal of Virology,* 85**,** 5125-5135.

AHN, J.-H., JANG, W.-J. & HAYWARD, G. S. 1999. The Human Cytomegalovirus IE2 and UL112-113 Proteins Accumulate in Viral DNA Replication Compartments That Initiate from the Periphery of Promyelocytic Leukemia Protein-Associated Nuclear Bodies (PODs or ND10). *Journal of Virology,* 73**,** 10458-10471.

AICHELER, R., WANG, E., TOMASEC, P., WILKINSON, G. & STANTON, R. 2013. Potential for Natural Killer Cell-Mediated Antibody-Dependent Cellular Cytotoxicity for Control of Human Cytomegalovirus. *Antibodies,* 2**,** 617.

AKRIGG, A., WILKINSON, G. W. & ORAM, J. D. 1985. The structure of the major immediate early gene of human cytomegalovirus strain AD169. *Virus Res,* 2**,** 107-21.

AKTER, P., CUNNINGHAM, C., MCSHARRY, B. P., DOLAN, A., ADDISON, C., DARGAN, D. J., HASSAN-WALKER, A. F., EMERY, V. C., GRIFFITHS, P. D., WILKINSON, G. W. & DAVISON, A. J. 2003. Two novel spliced genes in human cytomegalovirus. *J Gen Virol,* 84**,** 1117-22.

ALVAREZ-CASTELAO, B., RUIZ-RIVAS, C., CASTA, #XF1, O, J., #XE9 & G. 2012. A Critical Appraisal of Quantitative Studies of Protein Degradation in the Framework of Cellular Proteostasis. *Biochemistry Research International,* 2012**,** 11.

ALWINE, J. C. 2012. The Human Cytomegalovirus Assembly Compartment: A Masterpiece of Viral Manipulation of Cellular Processes That Facilitates Assembly and Egress. *PLOS Pathogens,* 8**,** e1002878.

ANDERHOLM, K. M., BIERLE, C. J. & SCHLEISS, M. R. 2016. Cytomegalovirus Vaccines: Current Status and Future Prospects. *Drugs,* 76**,** 1625-1645.

ANDERS, D. G., KACICA, M. A., PARI, G. & PUNTURIERI, S. M. 1992. Boundaries and structure of human cytomegalovirus oriLyt, a complex origin for lytic-phase DNA replication. *J Virol,* 66**,** 3373-84.

ANDERS, D. G., KERRY, J. A. & PARI, G. S. 2007. DNA synthesis and late viral gene expression. *In:* ARVIN, A., CAMPADELLI-FIUME, G., MOCARSKI, E., MOORE, P. S., ROIZMAN, B., WHITLEY, R. & YAMANISHI, K. (eds.) *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.* Cambridge: Cambridge University Press.

ARAMĂ, V., MIHAILESCU, R., RADULESCU, M., ARAMA, S. S., STREINU-CERCEL, A. & YOULE, M. 2014. Clinical relevance of the plasma load of cytomegalovirus in patients infected with HIV--a survival analysis. *J Med Virol,* 86**,** 1821-7.

ARANGO DUQUE, G. & DESCOTEAUX, A. 2014. Macrophage Cytokines: Involvement in Immunity and Infectious Diseases. *Frontiers in Immunology,* 5**,** 491.

ARNON, T. I., ACHDOUT, H., LEVI, O., MARKEL, G., SALEH, N., KATZ, G., GAZIT, R., GONEN-GROSS, T., HANNA, J., NAHARI, E., PORGADOR, A., HONIGMAN, A., PLACHTER, B., MEVORACH, D., WOLF, D. G. & MANDELBOIM, O. 2005. Inhibition of the NKp30 activating receptor by pp65 of human cytomegalovirus. *Nat Immunol,* 6**,** 515-23.

ASHIRU, O., BENNETT, N. J., BOYLE, L. H., THOMAS, M., TROWSDALE, J. & WILLS, M. R. 2009. NKG2D ligand MICA is retained in the cis-Golgi apparatus by human cytomegalovirus protein UL142. *J Virol,* 83**,** 12345-54.

AVERY, R. K., MARTY, F. M., STRASFELD, L., LEE, I., ARRIETA, A., CHOU, S., TATAROWICZ, W. & VILLANO, S. 2010. Oral maribavir for treatment of refractory or resistant cytomegalovirus infections in transplant recipients. *Transpl Infect Dis,* 12**,** 489-96.

AZEVEDO, L. S., PIERROTTI, L. C., ABDALA, E., COSTA, S. F., STRABELLI, T. M. V., CAMPOS, S. V., RAMOS, J. F., LATIF, A. Z. A., LITVINOV, N., MALUF, N. Z., FILHO, H. H. C., PANNUTI, C. S., LOPES, M. H., DOS SANTOS, V. A., DA CRUZ GOUVEIA LINARDI, C., YASUDA, M. A. S. & DE SOUSA MARQUES, H. H. 2015. Cytomegalovirus infection in transplant recipients. *Clinics,* 70**,** 515-523.

BALDANTI, F., REVELLO, M. G., PERCIVALLE, E., LABO, N. & GERNA, G. 2003. Genomes of the Endothelial Cell-Tropic Variant and the Parental Toledo Strain of Human Cytomegalovirus Are Highly Divergent. *J Med Virol***,** 76-81.

BALDICK, C. J., JR. & SHENK, T. 1996. Proteins associated with purified human cytomegalovirus particles. *J Virol,* 70**,** 6097-105.

BARBER, G. N. 2001. Host defense, viruses and apoptosis. *Cell Death And Differentiation,* 8**,** 113.

BARR, T. A., BROWN, S., RYAN, G., ZHAO, J. & GRAY, D. 2007. TLR-mediated stimulation of APC: Distinct cytokine responses of B cells and dendritic cells. *European Journal of Immunology,* 37**,** 3040-3053.

BARRY, P. A. & CHANG, W. L. W. 2007. Primate Betaherpesviruses. *In:* ARVIN, A., CAMPADELLI-FIUME, G., MOCARSKI, E., MOORE, P. S., ROIZMAN, B., WHITLEY, R. & YAMANISHI, K. (eds.) *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.* Cambridge: Cambridge University Press.

BATTISTA, M. C., BERGAMINI, G., BOCCUNI, M. C., CAMPANINI, F., RIPALTI, A. & LANDINI, M. P. 1999. Expression and Characterization of a Novel Structural Protein of Human Cytomegalovirus, pUL25. *Journal of Virology,* 73**,** 3800-3809.

BECHTEL, J. T. & SHENK, T. 2002. Human cytomegalovirus UL47 tegument protein functions after entry and before immediate-early gene expression. *J Virol,* 76**,** 1043-50.

BEN-SASSON, S. Z., HU-LI, J., QUIEL, J., CAUCHETAUX, S., RATNER, M., SHAPIRA, I., DINARELLO, C. A. & PAUL, W. E. 2009. IL-1 acts directly on CD4 T cells to enhance their antigen-driven expansion and differentiation. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 7119-7124.

BERNSTEIN, D. I., REAP, E. A., KATEN, K., WATSON, A., SMITH, K., NORBERG, P., OLMSTED, R. A., HOEPER, A., MORRIS, J., NEGRI, S., MAUGHAN, M. F. & CHULAY, J. D. 2009. Randomized, double-blind, Phase 1 trial of an alphavirus replicon vaccine for cytomegalovirus in CMV seronegative adult volunteers. *Vaccine,* 28**,** 484-93.

BERNSTEIN, D. I., SCHLEISS, M. R., BERENCSI, K., GONCZOL, E., DICKEY, M., KHOURY, P., CADOZ, M., MERIC, C., ZAHRADNIK, J., DULIEGE, A. M. & PLOTKIN, S. 2002. Effect of previous or simultaneous immunization with canarypox expressing cytomegalovirus (CMV) glycoprotein B (gB) on response to subunit gB vaccine plus MF59 in healthy CMV-seronegative adults. *J Infect Dis,* 185**,** 686-90.

BESOLD, K., FRANKENBERG, N., PEPPERL-KLINDWORTH, S., KUBALL, J., THEOBALD, M., HAHN, G. & PLACHTER, B. 2007. Processing and MHC class I presentation of human cytomegalovirus pp65-derived peptides persist despite gpUS2–11-mediated immune evasion. *Journal of General Virology,* 88**,** 1429-1439.

BHELLA, D., RIXON, F. J. & DARGAN, D. J. 2000. Cryomicroscopy of human cytomegalovirus virions reveals more densely packed genomic DNA than in herpes simplex virus type 111Edited by A. Klug. *Journal of Molecular Biology,* 295**,** 155-161.

BIRON, C. A., NGUYEN, K. B., PIEN, G. C., COUSENS, L. P. & SALAZAR-MATHER, T. P. 1999. Natural killer cells in antiviral defense: function and regulation by innate cytokines. *Annu Rev Immunol,* 17**,** 189-220.

BODAGHI, B., JONES, T. R., ZIPETO, D., VITA, C., SUN, L., LAURENT, L., ARENZANA-SEISDEDOS, F., VIRELIZIER, J. L. & MICHELSON, S. 1998. Chemokine sequestration by viral chemoreceptors as a novel viral escape strategy: withdrawal of chemokines from the environment of cytomegalovirus-infected cells. *J Exp Med,* 188**,** 855-66.

BODAGHI, B., SLOBBE-VAN DRUNEN, M. E., TOPILKO, A., PERRET, E., VOSSEN, R. C., VAN DAM-MIERAS, M. C., ZIPETO, D., VIRELIZIER, J. L., LEHOANG, P., BRUGGEMAN, C. A. & MICHELSON, S. 1999. Entry of human cytomegalovirus into retinal pigment epithelial and endothelial cells by endocytosis. *Invest Ophthalmol Vis Sci,* 40**,** 2598-607.

BOEHME, K. W., GUERRERO, M. & COMPTON, T. 2006. Human cytomegalovirus envelope glycoproteins B and H are necessary for TLR2 activation in permissive cells. *J Immunol,* 177**,** 7094-102.

BOEHME, K. W., SINGH, J., PERRY, S. T. & COMPTON, T. 2004. Human cytomegalovirus elicits a coordinated cellular antiviral response via envelope glycoprotein B. *J Virol,* 78**,** 1202-11.

BOGGIO, R. & CHIOCCA, S. 2006. Viruses and sumoylation: recent highlights. *Current Opinion in Microbiology,* 9**,** 430-436.

BOGNER, E., RADSAK, K. & STINSKI, M. F. 1998. The Gene Product of Human Cytomegalovirus Open Reading Frame UL56 Binds the pac Motif and Has Specific Nuclease Activity. *Journal of Virology,* 72**,** 2259-2264.

BONIFACINO, J. S. & TRAUB, L. M. 2003. Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annu Rev Biochem,* 72**,** 395-447.

BOPPANA, S. B. & BRITT, W. J. 2013. Synopsis of Clinical Aspects of Human Cytomegalovirus Disease. *In:* REDDEHASE, M. J. (ed.) *Cytomegaloviruses from Molecular Pathogenesis to Intervention.* Norfolk, UK: Caister Academic Press.

BOPPANA, S. B., FOWLER, K. B., PASS, R. F., RIVERA, L. B., BRADFORD, R. D., LAKEMAN, F. D. & BRITT, W. J. 2005. Congenital cytomegalovirus infection: association between virus burden in infancy and hearing loss. *J Pediatr,* 146**,** 817-23.

BRADLEY, A. J., LURAIN, N. S., GHAZAL, P., TRIVEDI, U., CUNNINGHAM, C., BALUCHOVA, K., GATHERER, D., WILKINSON, G. W. G., DARGAN, D. J. & DAVISON, A. J. 2009. High-throughput sequence analysis of variants of human cytomegalovirus strains Towne and AD169. *Journal of General Virology,* 90**,** 2375-2380.

BREGANTE, S., BERTILSON, S., TEDONE, E., VAN LINT, M. T., TRESPI, G., MORDINI, N., BERISSO, G., GUALANDI, F., LAMPARELLI, T., FIGARI, O., BENVENUTO, F., RAIOLA, A. M. & BACIGALUPO, A. 2000. Foscarnet prophylaxis of cytomegalovirus infections in

patients undergoing allogeneic bone marrow transplantation (BMT): a dose-finding study. *Bone Marrow Transplant,* 26**,** 23-9.

BRESNAHAN, W. A. & SHENK, T. E. 2000. UL82 virion protein activates expression of immediate early viral genes in human cytomegalovirus-infected cells. *Proceedings of the National Academy of Sciences,* 97**,** 14506-14511.

BRINKWORTH, J. F. & THORN, M. 2013. Vertebrate Immune System Evolution and Comparative Primate Immunity. *In:* PECHENKINA, J. F. B. A. K. (ed.) *Primates, Pathogens and Evolution.* New York: Springer.

BRITT, W. J. & BOPPANA, S. 2004. Human cytomegalovirus virion proteins. *Hum Immunol,* 65**,** 395-402.

BROERS, A. E., VAN DER HOLT, R., VAN ESSER, J. W., GRATAMA, J. W., HENZEN-LOGMANS, S., KUENEN-BOUMEESTER, V., LOWENBERG, B. & CORNELISSEN, J. J. 2000. Increased transplant-related morbidity and mortality in CMV-seropositive patients despite highly effective prevention of CMV disease after allogeneic T-cell-depleted stem cell transplantation. *Blood,* 95**,** 2240-5.

BRONZINI, M., LUGANINI, A., DELL'OSTE, V., DE ANDREA, M., LANDOLFO, S. & GRIBAUDO, G. 2012. The US16 Gene of Human Cytomegalovirus Is Required for Efficient Viral Infection of Endothelial and Epithelial Cells. *Journal of Virology,* 86**,** 6875-6888.

BROWNE, E. P. & SHENK, T. 2003. Human cytomegalovirus UL83-coded pp65 virion protein inhibits antiviral gene expression in infected cells. *Proceedings of the National Academy of Sciences,* 100**,** 11439-11444.

BRUCE, C., STONE, K., GULCICEK, E. & WILLIAMS, K. 2013. Proteomics and the analysis of proteomic data: 2013 overview of current protein-profiling technologies. *Curr Protoc Bioinformatics,* Chapter 13**,** Unit 13.21.

BRUNO, L., CORTESE, M., MONDA, G., GENTILE, M., CALO, S., SCHIAVETTI, F., ZEDDA, L., CATTANEO, E., PICCIOLI, D., SCHAEFER, M., NOTOMISTA, E., MAIONE, D., CARFI, A., MEROLA, M. & UEMATSU, Y. 2016. Human cytomegalovirus pUL10 interacts with leukocytes and impairs TCR-mediated T-cell activation. *Immunol Cell Biol,* 94**,** 849-860.

BUCHKOVICH, N. J., MAGUIRE, T. G., PATON, A. W., PATON, J. C. & ALWINE, J. C. 2009. The Endoplasmic Reticulum Chaperone BiP/GRP78 Is Important in the Structure and Function of the Human Cytomegalovirus Assembly Compartment. *Journal of Virology,* 83**,** 11421-11428.

BUCHKOVICH, N. J., MAGUIRE, T. G., YU, Y., PATON, A. W., PATON, J. C. & ALWINE, J. C. 2008. Human Cytomegalovirus Specifically Controls the Levels of the Endoplasmic Reticulum Chaperone BiP/GRP78, Which Is Required for Virion Assembly. *Journal of Virology,* 82**,** 31-39.

BUGHIO, F., ELLIOTT, D. A. & GOODRUM, F. 2013. An Endothelial Cell-Specific Requirement for the UL133-UL138 Locus of Human Cytomegalovirus for Efficient Virus Maturation. *Journal of Virology,* 87**,** 3062-3075.

BÜSCHER, N., PAULUS, C., NEVELS, M., TENZER, S. & PLACHTER, B. 2015. The proteome of human cytomegalovirus virions and dense bodies is conserved across different strains. *Med Microbiol Immunol,* 204**,** 285-93.

BUTCHER, S. J., AITKEN, J., MITCHELL, J., GOWEN, B. & DARGAN, D. J. 1998. Structure of the human cytomegalovirus B capsid by electron cryomicroscopy and image reconstruction. *J Struct Biol,* 124**,** 70-6.

CALÓ, S., CORTESE, M., CIFERRI, C., BRUNO, L., GERREIN, R., BENUCCI, B., MONDA, G., GENTILE, M., KESSLER, T., UEMATSU, Y., MAIONE, D., LILJA, A. E., CARFÍ, A. & MEROLA, M. 2016. The Human Cytomegalovirus UL116 Gene Encodes an Envelope Glycoprotein Forming a Complex with gH Independently from gL. *Journal of Virology,* 90**,** 4926-4938.

CANNON, M. J., HYDE, T. B. & SCHMID, D. S. 2011. Review of cytomegalovirus shedding in bodily fluids and relevance to congenital cytomegalovirus infection. *Reviews in medical virology,* 21**,** 240-255.

CANTRELL, S. R. & BRESNAHAN, W. A. 2006. Human cytomegalovirus (HCMV) UL82 gene product (pp71) relieves hDaxx-mediated repression of HCMV replication. *J Virol,* 80**,** 6188-91.

CAPOSIO, P., RIERA, L., HAHN, G., LANDOLFO, S. & GRIBAUDO, G. 2004. Evidence that the human cytomegalovirus 46-kDa UL72 protein is not an active dUTPase but a late protein dispensable for replication in fibroblasts. *Virology,* 325**,** 264-76.

CARRARA, G., SARAIVA, N., GUBSER, C., JOHNSON, B. F. & SMITH, G. L. 2012. Six-transmembrane topology for Golgi anti-apoptotic protein (GAAP) and Bax inhibitor 1 (BI-1) provides model for the transmembrane Bax inhibitor-containing motif (TMBIM) family. *J Biol Chem,* 287**,** 15896-905.

CASADO, J. L., ARRIZABALAGA, J., MONTES, M., MARTÍ-BELDA, P., TURAL, C., PINILLA, J., GUTIERREZ, C., PORTU, J., SCHUURMAN, R. & GROUP, K. A. F. T. S. C.-A. S. 1999. Incidence and risk factors for developing cytomegalovirus retinitis in HIV-infected patients receiving protease inhibitor therapy. *AIDS,* 13**,** 1497-1502.

CAVALETTO, N., LUGANINI, A. & GRIBAUDO, G. 2015. Inactivation of the Human Cytomegalovirus US20 Gene Hampers Productive Viral Replication in Endothelial Cells. *J Virol*.

CAYATTE, C., SCHNEIDER-OHRUM, K., WANG, Z., IRRINKI, A., NGUYEN, N., LU, J., NELSON, C., SERVAT, E., GEMMELL, L., CITKOWICZ, A., LIU, Y., HAYES, G., WOO, J., VAN NEST, G., JIN, H., DUKE, G. & MCCORMICK, A. L. 2013. Cytomegalovirus vaccine strain towne-derived dense bodies induce broad cellular immune responses and neutralizing antibodies that prevent infection of fibroblasts and epithelial cells. *J Virol,* 87**,** 11107-20.

CEBULLA, C. M., MILLER, D. M., KNIGHT, D. A., BRIGGS, B. R., MCGAUGHY, V. & SEDMAK, D. D. 2000. Cytomegalovirus induces sialyl Lewis(x) and Lewis(x) on human endothelial cells. *Transplantation,* 69**,** 1202-9.

CEPEDA, V., ESTEBAN, M. & FRAILE-RAMOS, A. 2010. Human cytomegalovirus final envelopment on membranes containing both trans-Golgi network and endosomal markers. *Cell Microbiol,* 12**,** 386-404.

CHA, T. A., TOM, E., KEMBLE, G. W., DUKE, G. M., MOCARSKI, E. S. & SPAETE, R. R. 1996. Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains. *J Virol,* 70.

CHALUPNY, N. J., REIN-WESTON, A., DOSCH, S. & COSMAN, D. 2006. Down-regulation of the NKG2D ligand MICA by the human cytomegalovirus glycoprotein UL142. *Biochem Biophys Res Commun,* 346**,** 175-81.

CHAMBERS, J., ANGULO, A., AMARATUNGA, D., GUO, H., JIANG, Y., WAN, J. S., BITTNER, A., FRUEH, K., JACKSON, M. R., PETERSON, P. A., ERLANDER, M. G. & GHAZAL, P. 1999. DNA microarrays of the complex human cytomegalovirus genome: Profiling kinetic class with drug sensitivity of viral gene expression. *Journal of Virology,* 73**,** 5757-5766.

CHANG, C. P., VESOLE, D. H., NELSON, J., OLDSTONE, M. B. & STINSKI, M. F. 1989. Identification and expression of a human cytomegalovirus early glycoprotein. *Journal of Virology,* 63**,** 3330-3337.

CHANG, W. L. W., BAUMGARTH, N., YU, D. & BARRY, P. A. 2004. Human Cytomegalovirus-Encoded Interleukin-10 Homolog Inhibits Maturation of Dendritic Cells and Alters Their Functionality. *Journal of Virology,* 78**,** 8720-8731.

CHAPMAN, T. L., HEIKEMAN, A. P. & BJORKMAN, P. J. 1999. The inhibitory receptor LIR-1 uses a common binding interaction to recognize class I MHC molecules and the viral homolog UL18. *Immunity,* 11**,** 603-13.

CHARPAK-AMIKAM, Y., KUBSCH, T., SEIDEL, E., OIKNINE-DJIAN, E., CAVALETTO, N., YAMIN, R., SCHMIEDEL, D., WOLF, D., GRIBAUDO, G., MESSERLE, M., CICIN-SAIN, L. & MANDELBOIM, O. 2017. Human cytomegalovirus escapes immune recognition by NK cells through the downregulation of B7-H6 by the viral genes US18 and US20. *Scientific Reports,* 7**,** 8661.

CHEE, M. S., BANKIER, A. T., BECK, S., BOHNI, R., BROWN, C. M., CERNY, R., HORSNELL, T., HUTCHISON, C. A., 3RD, KOUZARIDES, T., MARTIGNETTI, J. A. & ET AL. 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Curr Top Microbiol Immunol,* 154**,** 125-69.

CHEERAN, M. C. J., LOKENSGARD, J. R. & SCHLEISS, M. R. 2009. Neuropathogenesis of Congenital Cytomegalovirus Infection: Disease Mechanisms and Prospects for Intervention. *Clinical Microbiology Reviews,* 22**,** 99-126.

CHEN, D. H., JIANG, H., LEE, M., LIU, F. & ZHOU, Z. H. 1999. Three-Dimensional Visualization of Tegument/Capsid Interactions in the Intact Human Cytomegalovirus. *Virology,* 260**,** 10-16.

CHEN, H., BEARDSLEY, G. P. & COEN, D. M. 2014. Mechanism of ganciclovir-induced chain termination revealed by resistant viral polymerase mutants with reduced exonuclease activity. *Proceedings of the National Academy of Sciences of the United States of America,* 111**,** 17462-17467.

CHILD, S. J., HAKKI, M., DE NIRO, K. L. & GEBALLE, A. P. 2004. Evasion of Cellular Antiviral Responses by Human Cytomegalovirus TRS1 and IRS1. *Journal of Virology,* 78**,** 197-205.

CHOO, Y. S. & ZHANG, Z. 2009. Detection of Protein Ubiquitination. *Journal of Visualized Experiments : JoVE***,** 1293.

COCHRANE, D. 2009. *Characterisation of the Human Cytomegalovirus Immunomdulatory Gene UL141.* PhD thesos, Cardiff University.

COLBERG-POLEY, A. M. 1996. Functional Roles of Immediate Early Proteins Encoded by the Human Cytomegalovirus UL36-38, UL115-119, TRS1/IRS1 and US3 Loci. *Intervirology***,** 350-360.

COLBERG-POLEY, A. M. & WILLIAMSON, C. D. 2013. Intracellular Sorting and Trafficking of Cytomegalovirus Proteins during Permissive Infection. *In:* REDDEHASE, M. J. (ed.) *Cytomegaloviruses: From Molecular Pathogenesis to Intervention.* Norfolk, UK: Caister Academic Press.

COMPTON, T., NEPOMUCENO, R. R. & NOWLIN, D. M. 1992. Human cytomegalovirus penetrates host cells by pH-independent fusion at the cell surface. *Virology,* 191**,** 387-95.

CORRALES-AGUILAR, E., HOFFMANN, K. & HENGEL, H. 2014a. CMV-encoded Fcγ receptors: modulators at the interface of innate and adaptive immunity. *Seminars in Immunopathology,* 36**,** 627-640.

CORRALES-AGUILAR, E., TRILLING, M., HUNOLD, K., FIEDLER, M., LE, V. T. K., REINHARD, H., EHRHARDT, K., MERCÉ-MALDONADO, E., ALIYEV, E., ZIMMERMANN, A., JOHNSON, D. C. & HENGEL, H. 2014b. Human Cytomegalovirus Fcγ Binding Proteins gp34 and gp68 Antagonize Fcγ Receptors I, II and III. *PLoS Pathogens,* 10**,** e1004131.

CORTESE, M., CALO, S., D'AURIZIO, R., LILJA, A., PACCHIANI, N. & MEROLA, M. 2012. Recombinant human cytomegalovirus (HCMV) RL13 binds human immunoglobulin G Fc. *PLoS One,* 7**,** e50166.

COSMAN, D., MÜLLBERG, J., SUTHERLAND, C. L., CHIN, W., ARMITAGE, R., FANSLOW, W., KUBIN, M. & CHALUPNY, N. J. 2001. ULBPs, Novel MHC Class I Related Molecules,

Bind to CMV Glycoprotein UL16 and Stimulate NK Cytotoxicity through the NKG2D Receptor. *Immunity,* 14**,** 123-133.

CRAWFORD, L. B., KIM, J. H., COLLINS-MCMILLEN, D., LEE, B.-J., LANDAIS, I., HELD, C., NELSON, J. A., YUROCHKO, A. D. & CAPOSIO, P. 2018. Human Cytomegalovirus Encodes a Novel FLT3 Receptor Ligand Necessary for Hematopoietic Cell Differentiation and Viral Reactivation. *mBio,* 9**,** e00682-18.

DARGAN, D. J., DOUGLAS, E., CUNNINGHAM, C., JAMIESON, F., STANTON, R. J., BALUCHOVA, K., MCSHARRY, B. P., TOMASEC, P., EMERY, V. C., PERCIVALLE, E., SARASINI, A., GERNA, G., WILKINSON, G. W. G. & DAVISON, A. J. 2010. Sequential mutations associated with adaptation of human cytomegalovirus to growth in cell culture. *J Gen Virol,* 91**,** 1535-46.

DAS, S., ORTIZ, D. A., GURCZYNSKI, S. J., KHAN, F. & PELLETT, P. E. 2014. Identification of Human Cytomegalovirus Genes Important for Biogenesis of the Cytoplasmic Virion Assembly Complex. *Journal of Virology,* 88**,** 9086-9099.

DAS, S. & PELLETT, P. E. 2007. Members of the HCMV US12 family of predicted heptaspanning membrane proteins have unique intracellular distributions, including association with the cytoplasmic virion assembly complex. *Virology,* 361**,** 263-273.

DAS, S. & PELLETT, P. E. 2011. Spatial relationships between markers for secretory and endosomal machinery in human cytomegalovirus-infected cells versus those in uninfected cells. *J Virol***,** 5864-5879.

DAS, S., SKOMOROVSKA-PROKVOLIT, Y., WANG, F. Z. & PELLETT, P. E. 2006. Infection-dependent nuclear localization of US17, a member of the US12 family of human cytomegalovirus-encoded seven-transmembrane proteins. *Journal of Virology,* 80**,** 1191-1203.

DAS, S., VASANJI, A. & PELLETT, P. E. 2007. Three-dimensional structure of the human cytomegalovirus cytoplasmic virion assembly complex includes a reoriented secretory apparatus. *J Virol,* 81**,** 11861-9.

DAUBY, N., KUMMERT, C., LECOMTE, S., LIESNARD, C., DELFORGE, M. L., DONNER, C. & MARCHANT, A. 2014. Primary human cytomegalovirus infection induces the expansion of virus-specific activated and atypical memory B cells. *J Infect Dis,* 210**,** 1275-85.

DAVISON, A. J. 2007a. Comparative analysis of the genomes. *In:* ARVIN A, C.-F. G., MOCARSKI E, ET AL., (ed.) *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.* Cambridge: Cambridge University Press.

DAVISON, A. J. 2007b. Overview of classification. *In:* ARVIN, A., CAMPADELLI-FIUME, G., MOCARSKI, E., MOORE, P. S., ROIZMAN, B., WHITLEY, R. & YAMANISHI, K. (eds.) *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.* Cambridge University Press: Cambridge.

DAVISON, A. J. 2010. Herpesvirus systematics. *Vet Microbiol,* 143.

DAVISON, A. J. 2011. Evolution of sexually transmitted and sexually transmissible human herpesviruses. *Annals of the New York Academy of Sciences,* 1230**,** E37-E49.

DAVISON, A. J., AKTER, P., CUNNINGHAM, C., DOLAN, A., ADDISON, C., DARGAN, D. J., HASSAN-WALKER, A. F., EMERY, V. C., GRIFFITHS, P. D. & WILKINSON, G. W. G. 2003a. Homology between the human cytomegalovirus RL11 gene family and human adenovirus E3 genes. *Journal of General Virology,* 84**,** 657-663.

DAVISON, A. J. & BHELLA, D. 2007. Comparative genome and virion structure. *In:* ARVIN, A., CAMPADELLI-FIUME, G., MOCARSKI, E. & ET AL. (eds.) *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.* Cambridge: Cambridge University Press.

DAVISON, A. J., DARGAN, D. J. & STOW, N. D. 2002. Fundamental and accessory systems in herpesviruses. *Antiviral Res,* 56**,** 1-11.

DAVISON, A. J., DOLAN, A., AKTER, P., ADDISON, C., DARGAN, D. J., ALCENDOR, D. J., MCGEOCH, D. J. & HAYWARD, G. S. 2003b. The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. *J Gen Virol,* 84**,** 17-28.

DAVISON, A. J., HOLTON, M., DOLAN, A., DARGAN, D.J., GATHERER D., AND HAYWARD, G.S. 2013. Comparative Genomics of Primate Cytomegaloviruses. *In:* REDDEHASE, M. J. (ed.) *Cytomegaloviruses from Molecular Pathogenesis to Intervention.* Norfolk, UK.: Caister Academic Press.

DAVISON, A. J. & STOW, N. D. 2005. New Genes from Old: Redeployment of dUTPase by Herpesviruses. *Journal of Virology,* 79**,** 12880-12892.

DE CASTRO, E., SIGRIST, C. J., GATTIKER, A., BULLIARD, V., LANGENDIJK-GENEVAUX, P. S., GASTEIGER, E., BAIROCH, A. & HULO, N. 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res,* 34**,** W362-5.

DEAYTON, J. R., SABIN, C. A., JOHNSON, M. A., EMERY, V. C., WILSON, P. & GRIFFITHS, P. D. 2004. Importance of cytomegalovirus viraemia in risk of disease progression and death in HIV-infected patients receiving highly active antiretroviral therapy. *The Lancet,* 363**,** 2116-2121.

DECKERS, M., HOFMANN, J., KREUZER, K.-A., REINHARD, H., EDUBIO, A., HENGEL, H., VOIGT, S. & EHLERS, B. 2009. High genotypic diversity and a novel variant of human cytomegalovirus revealed by combined UL33/UL55 genotyping with broad-range PCR. *Virology Journal,* 6**,** 210-210.

DEFILIPPIS, V. R., ALVARADO, D., SALI, T., ROTHENBURG, S. & FRÜH, K. 2010. Human Cytomegalovirus Induces the Interferon Response via the DNA Sensor ZBP1. *Journal of Virology,* 84**,** 585-598.

DEMERITT, I. B., MILFORD, L. E. & YUROCHKO, A. D. 2004. Activation of the NF-kappaB pathway in human cytomegalovirus-infected cells is necessary for efficient transactivation of the major immediate-early promoter. *J Virol,* 78**,** 4498-507.

DEMMLER, G. J. 2006. Cytomegalovirus Infection: Back to the Future or No More Elephants? *Clin Infect Dis,* 43**,** 1152-1153.

DENG, J., LU, P. D., ZHANG, Y., SCHEUNER, D., KAUFMAN, R. J., SONENBERG, N., HARDING, H. P. & RON, D. 2004. Translational repression mediates activation of nuclear factor kappa B by phosphorylated translation initiation factor 2. *Mol Cell Biol,* 24**,** 10161-8.

DINKEL, H., VAN ROEY, K., MICHAEL, S., KUMAR, M., UYAR, B., ALTENBERG, B., MILCHEVSKAYA, V., SCHNEIDER, M., KUHN, H., BEHRENDT, A., DAHL, S. L., DAMERELL, V., DIEBEL, S., KALMAN, S., KLEIN, S., KNUDSEN, A. C., MADER, C., MERRILL, S., STAUDT, A., THIEL, V., WELTI, L., DAVEY, N. E., DIELLA, F. & GIBSON, T. J. 2016. ELM 2016--data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Res,* 44**,** D294-300.

DOBSON, L., REMÉNYI, I. & TUSNÁDY, G. E. 2015. CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic Acids Research,* 43**,** W408-W412.

DOLAN, A., CUNNINGHAM, C., HECTOR, R. D., HASSAN-WALKER, A. F., LEE, L., ADDISON, C., DARGAN, D. J., MCGEOCH, D. J., GATHERER, D., EMERY, V. C., GRIFFITHS, P. D., SINZGER, C., MCSHARRY, B. P., WILKINSON, G. W. & DAVISON, A. J. 2004. Genetic content of wild-type human cytomegalovirus. *J Gen Virol,* 85**,** 1301-12.

DUNN, C., CHALUPNY, N. J., SUTHERLAND, C. L., DOSCH, S., SIVAKUMAR, P. V., JOHNSON, D. C. & COSMAN, D. 2003a. Human Cytomegalovirus Glycoprotein UL16 Causes Intracellular Sequestration of NKG2D Ligands, Protecting Against Natural Killer Cell Cytotoxicity. *The Journal of Experimental Medicine,* 197**,** 1427-1439.

DUNN, W., CHOU, C., LI, H., HAI, R., PATTERSON, D., STOLC, V., ZHU, H. & LIU, F. 2003b. Functional profiling of a human cytomegalovirus genome. *Proc Natl Acad Sci U S A,* 100**,** 14223-8.

DUTTA, N., LASHMIT, P., YUAN, J., MEIER, J. & STINSKI, M. F. 2015. The Human Cytomegalovirus UL133-138 Gene Locus Attenuates the Lytic Viral Cycle in Fibroblasts. *PLOS ONE,* 10**,** e0120946.

E, X. & KOWALIK, T. F. 2014. The DNA Damage Response Induced by Infection with Human Cytomegalovirus and Other Viruses. *Viruses,* 6**,** 2155-2185.

EAGLE, R. A., TRAHERNE, J. A., HAIR, J. R., JAFFERJI, I. & TROWSDALE, J. 2009. ULBP6/RAET1L is an additional human NKG2D ligand. *Eur J Immunol,* 39**,** 3207-16.

EAGLE, R. A. & TROWSDALE, J. 2007. Promiscuity and the single receptor: NKG2D. *Nat Rev Immunol,* 7**,** 737-44.

EINSELE, H., REUSSER, P., BORNHÄUSER, M., KALHS, P., EHNINGER, G., HEBART, H., CHALANDON, Y., KRÖGER, N., HERTENSTEIN, B. & ROHDE, F. 2006. Oral valganciclovir leads to higher exposure to ganciclovir than intravenous ganciclovir in patients following allogeneic stem cell transplantation. *Blood,* 107**,** 3002-3008.

ELDE, N. C., CHILD, S. J., EICKBUSH, M. T., KITZMAN, J. O., ROGERS, K. S., SHENDURE, J., GEBALLE, A. P. & MALIK, H. S. 2012. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. *Cell,* 150**,** 831-41.

ELEK, S. D. & STERN, H. 1974. Development of a vaccine against mental retardation caused by cytomegalovirus infection in utero. *Lancet,* 1**,** 1-5.

EMERY, V. 2001. Investigation of CMV disease in immunocompromised patients. *Journal of Clinical Pathology,* 54**,** 84-88.

EMERY, V. C., COPE, A. V., BOWEN, E. F., GOR, D. & GRIFFITHS, P. D. 1999. The Dynamics of Human Cytomegalovirus Replication in Vivo. *The Journal of Experimental Medicine,* 190**,** 177-182.

EMERY, V. C., MILNE, R. S. B. & GRIFFITHS, P. D. 2013. Clinical Cytomegalovirus Research: Liver and Kidney Transplantation. *In:* REDDEHASE, M. J. (ed.) *Cytomegaloviruses from Molecular Pathogenesis to Intervention* Norfolk, UK: Caister Academic Press.

ENDERS, G., DAIMINGER, A., BADER, U., EXLER, S. & ENDERS, M. 2011. Intrauterine transmission and clinical outcome of 248 pregnancies with primary cytomegalovirus infection in relation to gestational age. *J Clin Virol,* 52**,** 244-6.

ENDERS, J. F., WELLER, T. H. & ROBBINS, F. C. 1949. Cultivation of the Lansing Strain of Poliomyelitis Virus in Cultures of Various Human Embryonic Tissues. *Science,* 109**,** 85-7.

ENGEL, P., PEREZ-CARMONA, N., ALBA, M. M., ROBERTSON, K., GHAZAL, P. & ANGULO, A. 2011. Human cytomegalovirus UL7, a homologue of the SLAM-family receptor CD229, impairs cytokine production. *Immunol Cell Biol,* 89**,** 753-766.

ERARD, V., GUTHRIE, K. A., SEO, S., SMITH, J., HUANG, M., CHIEN, J., FLOWERS, M. E. D., COREY, L. & BOECKH, M. 2015. Reduced Mortality of Cytomegalovirus Pneumonia After Hematopoietic Cell Transplantation Due to Antiviral Therapy and Changes in Transplantation Practices. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America,* 61**,** 31-39.

FEIRE, A. L., KOSS, H. & COMPTON, T. 2004. Cellular integrins function as entry receptors for human cytomegalovirus via a highly conserved disintegrin-like domain. *Proc Natl Acad Sci U S A,* 101**,** 15470-5.

FENG, X., SCHRÖER, J., YU, D. & SHENK, T. 2006. Human Cytomegalovirus pUS24 Is a Virion Protein That Functions Very Early in the Replication Cycle. *Journal of Virology,* 80**,** 8371-8378.

FIELDING, C. A., AICHELER, R., STANTON, R. J., WANG, E. C. Y., HAN, S., SEIRAFIAN, S., DAVIES, J., MCSHARRY, B. P., WEEKES, M. P., ANTROBUS, P. R., PROD'HOMME, V.,

BLANCHET, F. P., SUGRUE, D., CUFF, S., ROBERTS, D., DAVISON, A. J., LEHNER, P. J., WILKINSON, G. W. G. & TOMASEC, P. 2014. Two Novel Human Cytomegalovirus NK Cell Evasion Functions Target MICA for Lysosomal Degradation. *PLoS Pathogens,* 10.

FIELDING, C. A., WEEKES, M. P., NOBRE, L. V., RUCKOVA, E., WILKIE, G., PAULO, J. A., CHANG, C., SUAREZ, N., DAVIES, J., ANTROBUS, R., STANTON, R. J., AICHELER, R., NICHOLS, H., VOJTESEK, B., TROWSDALE, J., DAVISON, A. J., GYGI, S. P., TOMASEC, P., LEHNER, P. J. & WILKINSON, G. W. G. 2017. Control of immune ligands by members of a cytomegalovirus gene expansion suppresses natural killer cell activation *eLife,* 6**,** e22206.

FIELDS, B. N., KNIPE, D. M. & HOWLEY, P. M. 2013. *Fields Virology,* Philadelphia, Lippincott Williams & Wilkins.

FISH, K. N., SODERBERG-NAUCLER, C., MILLS, L. K., STENGLEIN, S. & NELSON, J. A. 1998. Human Cytomegalovirus Persistently Infects Aortic Endothelial Cells. *Journal of Virology,* 72**,** 5661-5668.

FOWLER, K. B., STAGNO, S. & PASS, R. F. 1993. Maternal age and congenital cytomegalovirus infection: screening of two diverse newborn populations, 1980-1990. *J Infect Dis,* 168**,** 552-6.

FOWLER, K. B., STAGNO, S. & PASS, R. F. 2003. Maternal immunity and prevention of congenital cytomegalovirus infection. *Jama,* 289**,** 1008-11.

FOWLER , K. B., STAGNO , S., PASS , R. F., BRITT , W. J., BOLL , T. J. & ALFORD , C. A. 1992. The Outcome of Congenital Cytomegalovirus Infection in Relation to Maternal Antibody Status. *New England Journal of Medicine,* 326**,** 663-667.

FREEZE, H. H. & KRANZ, C. 2010. Endoglycosidase and Glycoamidase Release of N-Linked Glycans. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.],* 0 17**,** 10.1002/0471142727.mb1713as89.

FREY, S. E., HARRISON, C., PASS, R. F., YANG, E., BOKEN, D., SEKULOVICH, R. E., PERCELL, S., IZU, A. E., HIRABAYASHI, S., BURKE, R. L. & DULIEGE, A. M. 1999. Effects of antigen dose and immunization regimens on antibody responses to a cytomegalovirus glycoprotein B subunit vaccine. *J Infect Dis,* 180**,** 1700-3.

FRIEDMAN, S. & FORD-JONES, E. L. 1999. Congenital cytomegalovirus infection – An update. *Paediatrics & Child Health,* 4**,** 35-38.

FU, Y. Z., SU, S., GAO, Y. Q., WANG, P. P., HUANG, Z. F., HU, M. M., LUO, W. W., LI, S., LUO, M. H., WANG, Y. Y. & SHU, H. B. 2017. Human Cytomegalovirus Tegument Protein UL82 Inhibits STING-Mediated Signaling to Evade Antiviral Immunity. *Cell Host Microbe,* 21**,** 231-243.

FURMAN, M. H., DEY, N., TORTORELLA, D. & PLOEGH, H. L. 2002. The Human Cytomegalovirus US10 Gene Product Delays Trafficking of Major Histocompatibility Complex Class I Molecules. *Journal of Virology,* 76**,** 11753-11756.

GABAEV, I., STEINBRUECK, L., POKOYSKI, C., PICH, A., STANTON, R. J., SCHWINZER, R., SCHULZ, T. F., JACOBS, R., MESSERLE, M. & KAY-FEDOROV, P. C. 2011. The Human Cytomegalovirus UL11 Protein Interacts with the Receptor Tyrosine Phosphatase CD45, Resulting in Functional Paralysis of T Cells. *Plos Pathogens,* 7.

GAO, Y., COLLETTI, K. & PARI, G. S. 2008. Identification of Human Cytomegalovirus UL84 Virus- and Cell-Encoded Binding Partners by Using Proteomics Analysis. *Journal of Virology,* 82**,** 96-104.

GARCÍA-GARCÍA, E. & ROSALES, C. 2000-2013. Adding Complexity to Phagocytic Signaling: Phagocytosis-Associated Cell Responses and Phagocytic Efficiency. *Madame Curie Bioscience Database [Internet].* Austin (TX): Landes Bioscience.

GASSER, S. & RAULET, D. H. 2006. Activation and self-tolerance of natural killer cells. *Immunological Reviews***,** 130-142.

GATHERER, D., SEIRAFIAN, S., CUNNINGHAM, C., HOLTON, M., DARGAN, D. J., BALUCHOVA, K., HECTOR, R. D., GALBRAITH, J., HERZYK, P., WILKINSON, G. W. G. & DAVISON, A. J. 2011. High-resolution Human Cytomegalovirus Transcriptome. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** 19755-19760.

GENINI, E., PERCIVALLE, E., SARASINI, A., REVELLO, M. G., BALDANTI, F. & GERNA, G. 2011. Serum antibody response to the gH/gL/pUL128-131 five-protein complex of human cytomegalovirus (HCMV) in primary and reactivated HCMV infections. *Journal of Clinical Virology,* 52**,** 113-118.

GERNA, G., REVELLO, M. G., BALDANTI, F., PERCIVALLE, E. & LILLERI, D. 2017. The pentameric complex of human Cytomegalovirus: cell tropism, virus dissemination, immune response and vaccine development. *Journal of General Virology,* 98**,** 2215-2234.

GERNA, G., SARASINI, A., PATRONE, M., PERCIVALLE, E., FIORINA, L., CAMPANINI, G., GALLINA, A., BALDANTI, F. & REVELLO, M. G. 2008. Human cytomegalovirus serum neutralizing antibodies block virus infection of endothelial/epithelial cells, but not fibroblasts, early during primary infection. *J Gen Virol,* 89**,** 853-65.

GIBSON, W. 1996. Structure and assembly of the virion. *Intervirology,* 39**,** 389-400.

GOLDMACHER, V. S. 2005. Cell death suppression by cytomegaloviruses. *Apoptosis,* 10**,** 251-65.

GOLDMACHER, V. S., BARTLE, L. M., SKALETSKAYA, A., DIONNE, C. A., KEDERSHA, N. L., VATER, C. A., HAN, J. W., LUTZ, R. J., WATANABE, S., CAHIR MCFARLAND, E. D., KIEFF, E. D., MOCARSKI, E. S. & CHITTENDEN, T. 1999. A cytomegalovirus-encoded mitochondria-localized inhibitor of apoptosis structurally unrelated to Bcl-2. *Proc Natl Acad Sci U S A,* 96**,** 12536-41.

GOODRUM, F. 2016. Human Cytomegalovirus Latency: Approaching the Gordian Knot. *Annual review of virology,* 3**,** 333-357.

GOODRUM, F. D., JORDAN, C. T., HIGH, K. & SHENK, T. 2002. Human cytomegalovirus gene expression during infection of primary hematopoietic progenitor cells: a model for latency. *Proc Natl Acad Sci U S A,* 99.

GRATTAN, M. T., MORENO-CABRAL, C. E., STARNES, V. A., OYER, P. E., STINSON, E. B. & SHUMWAY, N. E. 1989. Cytomegalovirus infection is associated with cardiac allograft rejection and atherosclerosis. *JAMA,* 261**,** 3561-3566.

GREFTE, A., VAN DER GIESSEN, M., VAN SON, W. & THE, T. H. 1993. Circulating cytomegalovirus (CMV)-infected endothelial cells in patients with an active CMV infection. *J Infect Dis,* 167**,** 270-7.

GREY, F., MEYERS, H., WHITE, E. A., SPECTOR, D. H. & NELSON, J. 2007. A Human Cytomegalovirus-Encoded microRNA Regulates Expression of Multiple Viral Genes Involved in Replication. *PLOS Pathogens,* 3**,** e163.

GRIFFIN, C., WANG, E. C. Y., MCSHARRY, B. P., RICKARDS, C., BROWNE, H., WILKINSON, G. W. G. & TOMASEC, P. 2005. Characterization of a highly glycosylated form of the human cytomegalovirus HLA class I homologue gpUL18. *The Journal of general virology,* 86**,** 2999-3008.

GRIFFIN, G. K., NEWTON, G., TARRIO, M. L., BU, D. X., MAGANTO-GARCIA, E., AZCUTIA, V., ALCAIDE, P., GRABIE, N., LUSCINSKAS, F. W., CROCE, K. J. & LICHTMAN, A. H. 2012. IL-17 and TNF-alpha sustain neutrophil recruitment during inflammation through synergistic effects on endothelial activation. *J Immunol,* 188**,** 6287-99.

GRIFFITHS, P. D., STANTON, A., MCCARRELL, E., SMITH, C., OSMAN, M., HARBER, M., DAVENPORT, A., JONES, G., WHEELER, D. C., O'BEIRNE, J., THORBURN, D., PATCH, D., ATKINSON, C. E., PICHON, S., SWENY, P., LANZMAN, M., WOODFORD, E., ROTHWELL, E., OLD, N., KINYANJUI, R., HAQUE, T., ATABANI, S., LUCK, S., PRIDEAUX, S., MILNE, R. S. B., EMERY, V. C. & BURROUGHS, A. K. 2011.

Cytomegalovirus glycoprotein-B vaccine with MF59 adjuvant in transplant recipients: a phase 2 randomised placebo-controlled trial. *Lancet,* 377**,** 1256-1263.

GRINDE, B. 2013. Herpesviruses: latency and reactivation – viral strategies and host response. *Journal of Oral Microbiology,* 5**,** 10.3402/jom.v5i0.22766.

GRINDE, B. & SEGLEN, P. O. 1980. Differential effects of proteinase inhibitors and amines on the lysosomal and non-lysosomal pathways of protein degradation in isolated rat hepatocytes. *Biochimica et Biophysica Acta (BBA) - General Subjects,* 632**,** 73-86.

GROSSE, S. D., ORTEGO-SANCHEZ, I. R., BIALEK, S. R. & DOLLARD, S. C. 2013. The Economic Impact of Congenital CMV Infection: Methods and Estimates. *In:* REDDEHASE, M. J. (ed.) *Cytomegaloviruses: From Molecular Pathogenesis to Intervention* Norfolk, UK: Caister Academic Press.

GROVES, I. J., REEVES, M. B. & SINCLAIR, J. H. 2009. Lytic infection of permissive cells with human cytomegalovirus is regulated by an intrinsic 'pre-immediate-early' repression of viral gene expression mediated by histone post-translational modification. *Journal of General Virology,* 90**,** 2364-2374.

GUETTA, E., SCARPATI, E. M. & DICORLETO, P. E. 2001. Effect of cytomegalovirus immediate early gene products on endothelial cell gene activity. *Cardiovasc Res,* 50**,** 538-46.

GUO, H., SHEN, S., WANG, L. & DENG, H. 2010. Role of tegument proteins in herpesvirus assembly and egress. *Protein Cell,* 1**,** 987-98.

GUO, Y. W., AND HUANG, E.S. 1993. Characterization of a Structurally Tricistronic Gene of Human Cytomegalovirus Composed of Us18, Us19, and Us20. *Journal of Virology,* 67**,** 2043-2054.

GUPTA, R., JUNG, E. & BRUNAK, S. 2004. *Prediction of N-glycosylation sites in human proteins*.

GURCZYNSKI, S. J., DAS, S. & PELLETT, P. E. 2014. Deletion of the Human Cytomegalovirus US17 Gene Increases the Ratio of Genomes per Infectious Unit and Alters Regulation of Immune and Endoplasmic Reticulum Stress Response Genes at Early and Late Times after Infection. *Journal of Virology,* 88**,** 2168-2182.

HAHN, G., REVELLO, M. G., PATRONE, M., PERCIVALLE, E., CAMPANINI, G., SARASINI, A., WAGNER, M., GALLINA, A., MILANESI, G., KOSZINOWSKI, U., BALDANTI, F. & GERNA, G. 2004. Human cytomegalovirus UL131-128 genes are indispensable for virus growth in endothelial cells and virus transfer to leukocytes. *J Virol,* 78**,** 10023-33.

HAI, R., CHU, A., LI, H., UMAMOTO, S., RIDER, P. & LIU, F. 2006. Infection of human cytomegalovirus in cultured human gingival tissue. *Virology Journal,* 3**,** 84-84.

HALENIUS, A., HAUKA, S., DÖLKEN, L., STINDT, J., REINHARD, H., WIEK, C., HANENBERG, H., KOSZINOWSKI, U. H., MOMBURG, F. & HENGEL, H. 2011. Human Cytomegalovirus Disrupts the Major Histocompatibility Complex Class I Peptide-Loading Complex and Inhibits Tapasin Gene Transcription. *Journal of Virology,* 85**,** 3473-3485.

HANSEN, S. G., STRELOW, L. I., FRANCHI, D. C., ANDERS, D. G. & WONG, S. W. 2003. Complete Sequence and Genomic Analysis of Rhesus Cytomegalovirus. *Journal of Virology,* 77**,** 6620-6636.

HASSAN, J., DOOLEY, S. & HALL, W. 2007. Immunological response to cytomegalovirus in congenitally infected neonates. *Clinical and Experimental Immunology,* 147**,** 465-471.

HEGDE, N. R., CHEVALIER, M. S., WISNER, T. W., DENTON, M. C., SHIRE, K., FRAPPIER, L. & JOHNSON, D. C. 2006. The Role of BiP in Endoplasmic Reticulum-associated Degradation of Major Histocompatibility Complex Class I Heavy Chain Induced by Cytomegalovirus Proteins. *Journal of Biological Chemistry,* 281**,** 20910-20919.

HELENIUS, A. 1994. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Molecular Biology of the Cell,* 5**,** 253-265.

HENSEL, G., MEYER, H., GARTNER, S., BRAND, G. & KERN, H. F. 1995. Nuclear localization of the human cytomegalovirus tegument protein pp150 (ppUL32). *J Gen Virol,* 76 (Pt 7)**,** 1591-601.

HERTEL, L. & MOCARSKI, E. S. 2004. Global analysis of host cell gene expression late during cytomegalovirus infection reveals extensive dysregulation of cell cycle gene expression and induction of Pseudomitosis independent of US28 function. *J Virol,* 78**,** 11988-2011.

HODSON, E. M., CRAIG, J. C., STRIPPOLI, G. F. & WEBSTER, A. C. 2008. Antiviral medications for preventing cytomegalovirus disease in solid organ transplant recipients. *Cochrane Database Syst Rev***,** Cd003774.

HOFMANN, K. & STOFFEL, W. 1993. TMbase - A database of membrane spanning proteins segments. *Biol. Chem. Hoppe-Seyler***,** 166.

HOLCIK, M. & SONENBERG, N. 2005. Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol,* 6**,** 318-27.

HOLLINSHEAD, M., JOHNS, H. L., SAYERS, C. L., GONZALEZ-LOPEZ, C., SMITH, G. L. & ELLIOTT, G. 2012. Endocytic tubules regulated by Rab GTPases 5 and 11 are used for envelopment of herpes simplex virus. *Embo j,* 31**,** 4204-20.

HOLZERLANDT, R., ORENGO, C., KELLAM, P. & ALBA, M. M. 2002. Identification of new herpesvirus gene homologs in the human genome. *Genome Res,* 12**,** 1739-48.

HOMMAN-LOUDIYI, M., HULTENBY, K., BRITT, W. & SODERBERG-NAUCLER, C. 2003. Envelopment of human cytomegalovirus occurs by budding into Golgi-derived vacuole compartments positive for gB, Rab 3, trans-golgi network 46, and mannosidase II. *J Virol,* 77**,** 3191-203.

HSU, J.-L., VAN DEN BOOMEN, D. J. H., TOMASEC, P., WEEKES, M. P., ANTROBUS, R., STANTON, R. J., RUCKOVA, E., SUGRUE, D., WILKIE, G. S., DAVISON, A. J., WILKINSON, G. W. G. & LEHNER, P. J. 2015. Plasma Membrane Profiling Defines an Expanded Class of Cell Surface Proteins Selectively Targeted for Degradation by HCMV US2 in Cooperation with UL141. *PLOS Pathogens,* 11**,** e1004811.

HUANG, E.-S. & JOHNSON, R. A. 2000. Human cytomegalovirus - no longer just a DNA virus. *Nat Med,* 6**,** 863-864.

HUBER, M. T., TOMAZIN, R., WISNER, T., BONAME, J. & JOHNSON, D. C. 2002. Human Cytomegalovirus US7, US8, US9, and US10 Are Cytoplasmic Glycoproteins, Not Found at Cell Surfaces, and US9 Does Not Mediate Cell-to-Cell Spread. *Journal of Virology,* 76**,** 5748-5758.

ICTV, T. R. 2011. Virus Taxonomy: The Classification and Nomenclature of Viruses. Order: Herpesviridales. *The 9th Report of the International Committee of the Taxonomy of Viruses.* Oxford: Elsevier.

IRMIERE, A. & GIBSON, W. 1983. Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. *Virology,* 130**,** 118-33.

ISAACSON, M. K. & COMPTON, T. 2009. Human cytomegalovirus glycoprotein B is required for virus entry and cell-to-cell spread but not for virion attachment, assembly, or egress. *J Virol,* 83**,** 3891-903.

ISHIDA, J. H., PATEL, A., MEHTA, A. K., GATAULT, P., MCBRIDE, J. M., BURGESS, T., DERBY, M. A., SNYDMAN, D. R., EMU, B., FEIERBACH, B., FOUTS, A. E., MAIA, M., DENG, R., ROSENBERGER, C. M., GENNARO, L. A., STRIANO, N. S., LIAO, X. C. & TAVEL, J. A. 2017. Phase 2 Randomized, Double-Blind, Placebo-Controlled Trial of RG7667, a Combination Monoclonal Antibody, for Prevention of Cytomegalovirus Infection in High-Risk Kidney Transplant Recipients. *Antimicrob Agents Chemother,* 61.

ISKENDERIAN, A. C., HUANG, L., REILLY, A., STENBERG, R. M. & ANDERS, D. G. 1996. Four of eleven loci required for transient complementation of human cytomegalovirus DNA

replication cooperate to activate expression of replication genes. *J Virol,* 70**,** 383-92.

ISLER, J. A., SKALET, A. H. & ALWINE, J. C. 2005. Human cytomegalovirus infection activates and regulates the unfolded protein response. *J Virol,* 79**,** 6890-9.

JANEWAY, C. A. J., TRAVERS, P., WALPORT, M. & SHLOMCHIK, M. J. 2001. Principles of innate and adaptive immunity. *Immunobiology: The Immune System in Health and Disease.* . 5th ed. New York: Garland Science.

JARVIS, M. A. & NELSON, J. A. 2002. Human cytomegalovirus persistence and latency in endothelial cells and macrophages. *Curr Opin Microbiol,* 5**,** 403-7.

JARVIS, M. A. & NELSON, J. A. 2007. Human Cytomegalovirus Tropism for Endothelial Cells: Not All Endothelial Cells Are Created Equal. *Journal of Virology,* 81**,** 2095-2101.

JEAN BELTRAN, P. M. & CRISTEA, I. M. 2014. The life cycle and pathogenesis of human cytomegalovirus infection: lessons from proteomics. *Expert review of proteomics,* 11**,** 697-711.

JEAN BELTRAN, P. M., MATHIAS, R. A. & CRISTEA, I. M. 2016. A portrait of the human organelle proteome in space and time during cytomegalovirus infection. *Cell systems,* 3**,** 361-373.e6.

JONES, T. R. & MUZITHRAS, V. P. 1991. Fine mapping of transcripts expressed from the US6 gene family of human cytomegalovirus strain AD169. *J Virol,* 65**,** 2024-36.

JONES, T. R. & SUN, L. 1997. Human cytomegalovirus US2 destabilizes major histocompatibility complex class I heavy chains. *J Virol,* 71**,** 2970-9.

KAISHO, T. & AKIRA, S. 2006. Toll-like receptor function and signaling. *J Allergy Clin Immunol,* 117**,** 979-87; quiz 988.

KALEJTA, R. F. 2008. Tegument Proteins of Human Cytomegalovirus. *Microbiology and Molecular Biology Reviews : MMBR,* 72**,** 249-265.

KALL, L., KROGH, A. & SONNHAMMER, E. L. 2004. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol,* 338**,** 1027-36.

KARAVELLAS, M. P., PLUMMER, D. J., MACDONALD, J. C., TORRIANI, F. J., SHUFELT, C. L., AZEN, S. P. & FREEMAN, W. R. 1999. Incidence of immune recovery vitritis in cytomegalovirus retinitis patients following institution of successful highly active antiretroviral therapy. *J Infect Dis,* 179**,** 697-700.

KASPARI, M., TAVALAI, N., STAMMINGER, T., ZIMMERMANN, A., SCHILF, R. & BOGNER, E. 2008. Proteasome inhibitor MG132 blocks viral DNA replication and assembly of human cytomegalovirus. *FEBS Lett. ,* 582**,** 666-72.

KATZE, M. G., HE, Y. & GALE, M., JR. 2002. Viruses and interferon: a fight for supremacy. *Nat Rev Immunol,* 2**,** 675-87.

KELLY, C., VAN DRIEL, R. & WILKINSON, G. W. G. 1995. Disruption of PML-associated nuclear bodies during human cytomegalovirus infection. *Journal of General Virology,* 76**,** 2887-2893.

KENNESON, A. & CANNON, M. J. 2007. Review and meta-analysis of the epidemiology of congenital cytomegalovirus (CMV) infection. *Rev Med Virol,* 17**,** 253-76.

KEYES, L. R., HARGETT, D., SOLAND, M., BEGO, M. G., ROSSETTO, C. C., ALMEIDA-PORADA, G. & ST JEOR, S. 2012. HCMV protein LUNA is required for viral reactivation from latently infected primary CD14(+) cells. *PLoS One,* 7**,** e52827.

KHAIRALLAH, C., DÉCHANET-MERVILLE, J. & CAPONE, M. 2017. γδ T Cell-Mediated Immunity to Cytomegalovirus Infection. *Frontiers in Immunology,* 8**,** 105.

KHARFAN-DABAJA, M. A., BOECKH, M., WILCK, M. B., LANGSTON, A. A., CHU, A. H., WLOCH, M. K., GUTERWILL, D. F., SMITH, L. R., ROLLAND, A. P. & KENNEY, R. T. 2012. A novel therapeutic cytomegalovirus DNA vaccine in allogeneic haemopoietic stem-cell transplantation: a randomised, double-blind, placebo-controlled, phase 2 trial. *The Lancet Infectious Diseases,* 12**,** 290-299.

KLEDAL, T. N., ROSENKILDE, M. M. & SCHWARTZ, T. W. 1998. Selective recognition of the membrane-bound CX3C chemokine, fractalkine, by the human cytomegalovirus-encoded broad-spectrum receptor US28. *FEBS Lett,* 441**,** 209-14.

KOBILER, O., DRAYMAN, N., BUTIN-ISRAELI, V. & OPPENHEIM, A. 2012. Virus strategies for passing the nuclear envelope barrier. *Nucleus,* 3**,** 526-539.

KOMINAMI, K., NAKABAYASHI, J., NAGAI, T., TSUJIMURA, Y., CHIBA, K., KIMURA, H., MIYAWAKI, A., SAWASAKI, T., YOKOTA, H., MANABE, N. & SAKAMAKI, K. 2012. The molecular mechanism of apoptosis upon caspase-8 activation: Quantitative experimental validation of a mathematical model. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1823**,** 1825-1840.

KOTENKO, S. V., SACCANI, S., IZOTOVA, L. S., MIROCHNITCHENKO, O. V. & PESTKA, S. 2000. Human cytomegalovirus harbors its own unique IL-10 homolog (cmvIL-10). *Proc Natl Acad Sci U S A,* 97**,** 1695-700.

KRZYZANIAK, M., MACH, M. & BRITT, W. J. 2007. The Cytoplasmic Tail of Glycoprotein M (gpUL100) Expresses Trafficking Signals Required for Human Cytomegalovirus Assembly and Replication. *Journal of Virology,* 81**,** 10316-10328.

LA ROSA, C. & DIAMOND, D. J. 2012. The immune response to human CMV. *Future virology,* 7**,** 279-293.

LA ROSA, C., LIMAYE, A. P., KRISHNAN, A., LONGMATE, J. & DIAMOND, D. J. 2007. Longitudinal Assessment of Cytomegalovirus (CMV)—Specific Immune Responses in Liver Transplant Recipients at High Risk for Late CMV Disease. *The Journal of Infectious Diseases,* 195**,** 633-644.

LANDINI, M. P., SEVERI, B., FURLINI, G. & BADIALI DE GIORGI, L. 1987. Human cytomegalovirus structural components: intracellular and intraviral localization of p28 and p65-69 by immunoelectron microscopy. *Virus Res,* 8**,** 15-23.

LANIER, L. L. 2008. Up on the tightrope: natural killer cell activation and inhibition. *Nature immunology,* 9**,** 495-502.

LANIER, L. L., CORLISS, B. C., WU, J., LEONG, C. & PHILLIPS, J. H. 1998. Immunoreceptor DAP12 bearing a tyrosine-based activation motif is involved in activating NK cells. *Nature,* 391**,** 703-7.

LANZIERI, T. M., DOLLARD, S. C., BIALEK, S. R. & GROSSE, S. D. 2014. Systematic review of the birth prevalence of congenital cytomegalovirus infection in developing countries. *International Journal of Infectious Diseases,* 22**,** 44-48.

LEE, K. Y., YOO, B.-W., AHN, S. S., BAE, W. H., LEE, H., JUNG, S. M., LEE, S.-W., PARK, Y.-B. & SONG, J. J. 2017. Predictors of mortality in autoimmune disease patients with concurrent cytomegalovirus infections detected by quantitative real-time PCR. *PLOS ONE,* 12**,** e0181590.

LEE, S. H., KALEJTA, R. F., KERRY, J., SEMMES, O. J., O'CONNOR, C. M., KHAN, Z., GARCIA, B. A., SHENK, T. & MURPHY, E. 2012. BclAF1 restriction factor is neutralized by proteasomal degradation and microRNA repression during human cytomegalovirus infection. *Proceedings of the National Academy of Sciences,* 109**,** 9575-9580.

LEHNER, P. J., KARTTUNEN, J. T., WILKINSON, G. W. & CRESSWELL, P. 1997. The human cytomegalovirus US6 glycoprotein inhibits transporter associated with antigen processing-dependent peptide translocation. *Proc Natl Acad Sci U S A,* 94**,** 6904-9.

LEON, R. P., HEDLUND, T., MEECH, S. J., LI, S., SCHAACK, J., HUNGER, S. P., DUKE, R. C. & DEGREGORI, J. 1998. Adenoviral-mediated gene transfer in lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 13159-13164.

LESNIEWSKI, M., DAS, S., SKOMOROVSKA-PROKVOLIT, Y., WANG, F.-Z. & PELLETT, P. E. 2006. Primate cytomegalovirus US12 gene family: A distinct and diverse clade of seven-transmembrane proteins. *Virology,* 354**,** 286-298.

LI, G., NGUYEN, C. C., RYCKMAN, B. J., BRITT, W. J. & KAMIL, J. P. 2015. A viral regulator of glycoprotein complexes contributes to human cytomegalovirus cell tropism. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** 4471-4476.

LI, T., CHEN, J. & CRISTEA, I. M. 2013. Human cytomegalovirus tegument protein pUL83 inhibits IFI16-mediated DNA sensing for immune evasion. *Cell host & microbe,* 14**,** 10.1016/j.chom.2013.10.007.

LIEBERMAN, J. 2010. Granzyme A activates another way to die. *Immunological reviews,* 235**,** 93-104.

LILLERI, D., KABANOVA, A., REVELLO, M. G., PERCIVALLE, E., SARASINI, A., GENINI, E., SALLUSTO, F., LANZAVECCHIA, A., CORTI, D. & GERNA, G. 2013. Fetal human cytomegalovirus transmission correlates with delayed maternal antibodies to gH/gL/pUL128-130-131 complex during primary infection. *PLoS One,* 8**,** e59863.

LILLEY, B. N., PLOEGH, H. L. & TIRABASSI, R. S. 2001. Human Cytomegalovirus Open Reading Frame TRL11/IRL11 Encodes an Immunoglobulin G Fc-Binding Protein. *Journal of Virology,* 75**,** 11218-11221.

LIMAYE, A. P., COREY, L., KOELLE, D. M., DAVIS, C. L. & BOECKH, M. 2000. Emergence of ganciclovir-resistant cytomegalovirus disease among recipients of solid-organ transplants. *The Lancet,* 356**,** 645-649.

LISAK, D. A., SCHACHT, T., ENDERS, V., HABICHT, J., KIVILUOTO, S., SCHNEIDER, J., HENKE, N., BULTYNCK, G. & METHNER, A. 2015. The transmembrane Bax inhibitor motif (TMBIM) containing protein family: Tissue expression, intracellular localization and effects on the ER CA2+-filling state. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1853**,** 2104-2114.

LIU, F. & ZHOU, Z. H. 2007. Comparative virion structures of human herpesviruses. *In:* ARVIN, A., CAMPADELLI-FIUME, G., MOCARSKI, E., ET AL., EDITORS. (ed.) *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.* Cambridge: Cambridge University Press.

LIU, Y. & BIEGALKE, B. J. 2002. The Human Cytomegalovirus UL35 Gene Encodes Two Proteins with Different Functions. *Journal of Virology,* 76**,** 2460-2468.

LJUNGGREN, H. G. & KARRE, K. 1986. Experimental strategies and interpretations in the analysis of changes in MHC gene expression during tumour progression. Opposing influences of T cell and natural killer mediated resistance? *J Immunogenet,* 13**,** 141-51.

LJUNGMAN, P., DELILIERS, G. L., PLATZBECKER, U., MATTHES-MARTIN, S., BACIGALUPO, A., EINSELE, H., ULLMANN, J., MUSSO, M., TRENSCHEL, R., RIBAUD, P., BORNHAUSER, M., CESARO, S., CROOKS, B., DEKKER, A., GRATECOS, N., KLINGEBIEL, T., TAGLIAFERRI, E., ULLMANN, A. J., WACKER, P. & CORDONNIER, C. 2001. Cidofovir for cytomegalovirus infection and disease in allogeneic stem cell transplant recipients. The Infectious Diseases Working Party of the European Group for Blood and Marrow Transplantation. *Blood,* 97**,** 388-92.

LJUNGMAN, P., ENGELHARD, D., LINK, H., BIRON, P., BRANDT, L., BRUNET, S., CORDONNIER, C., DEBUSSCHER, L., DE LAURENZI, A., KOLB, H. J. & ET AL. 1992. Treatment of interstitial pneumonitis due to cytomegalovirus with ganciclovir and intravenous immune globulin: experience of European Bone Marrow Transplant Group. *Clin Infect Dis,* 14**,** 831-5.

LOOMIS, R. J., LILJA, A. E., MONROE, J., BALABANIS, K. A., BRITO, L. A., PALLADINO, G., FRANTI, M., MANDL, C. W., BARNETT, S. W. & MASON, P. W. 2013. Vectored co-delivery of human cytomegalovirus gH and gL proteins elicits potent complement-independent neutralizing antibodies. *Vaccine,* 31**,** 919-926.

LU, Y., MA, Y., LIU, Z., HAN, L., GAO, S., ZHENG, B., LIU, C., QI, Y., SUN, Z., HUANG, Y. & RUAN, Q. 2016. A cluster of 3' coterminal transcripts from US12-US17 locus of human cytomegalovirus. *Virus Genes*.

LUGANINI, A., CAVALETTO, N., RAIMONDO, S., GEUNA, S. & GRIBAUDO, G. 2017. Loss of the Human Cytomegalovirus US16 Protein Abrogates Virus Entry into Endothelial and Epithelial Cells by Reducing the Virion Content of the Pentamer. *J Virol*.

LUO, M. H., HANNEMANN, H., KULKARNI, A. S., SCHWARTZ, P. H., O'DOWD, J. M. & FORTUNATO, E. A. 2010. Human cytomegalovirus infection causes premature and abnormal differentiation of human neural progenitor cells. *J Virol,* 84**,** 3528-41.

LÜTTICHAU, H. R. 2010. The Cytomegalovirus UL146 Gene Product vCXCL1 Targets Both CXCR1 and CXCR2 as an Agonist. *The Journal of Biological Chemistry,* 285**,** 9137-9146.

MA, Y., GAO, S., WANG, L., WANG, N., LI, M., ZHENG, B., QI, Y., SUN, Z., LIU, W. & RUAN, Q. 2013. Analysis and mapping of a 3' coterminal transcription unit derived from human cytomegalovirus open reading frames UL30–UL32. *Virology Journal,* 10**,** 65.

MACH, M., KROPFF, B., DAL MONTE, P. & BRITT, W. 2000. Complex Formation by Human Cytomegalovirus Glycoproteins M (gpUL100) and N (gpUL73). *Journal of Virology,* 74**,** 11881-11892.

MACH, M., OSINSKI, K., KROPFF, B., SCHLOETZER-SCHREHARDT, U., KRZYZANIAK, M. & BRITT, W. 2007. The Carboxy-Terminal Domain of Glycoprotein N of Human Cytomegalovirus Is Required for Virion Morphogenesis. *Journal of Virology,* 81**,** 5212-5224.

MACIEJEWSKI, J. P., BRUENING, E. E., DONAHUE, R. E., MOCARSKI, E. S., YOUNG, N. S. & ST JEOR, S. C. 1992. Infection of hematopoietic progenitor cells by human cytomegalovirus. *Blood,* 80**,** 170-8.

MANICKLAL, S., EMERY, V. C., LAZZAROTTO, T., BOPPANA, S. B. & GUPTA, R. K. 2013. The "silent" global burden of congenital cytomegalovirus. *Clin Microbiol Rev,* 26**,** 86-102.

MARSCHALL, M., MARZI, A., AUS DEM SIEPEN, P., JOCHMANN, R., KALMER, M., AUEROCHS, S., LISCHKA, P., LEIS, M. & STAMMINGER, T. 2005. Cellular p32 recruits cytomegalovirus kinase pUL97 to redistribute the nuclear lamina. *J Biol Chem,* 280**,** 33357-67.

MARTY, F. M., LJUNGMAN, P., CHEMALY, R. F., MAERTENS, J., DADWAL, S. S., DUARTE, R. F., HAIDER, S., ULLMANN, A. J., KATAYAMA, Y., BROWN, J., MULLANE, K. M., BOECKH, M., BLUMBERG, E. A., EINSELE, H., SNYDMAN, D. R., KANDA, Y., DINUBILE, M. J., TEAL, V. L., WAN, H., MURATA, Y., KARTSONIS, N. A., LEAVITT, R. Y. & BADSHAH, C. 2017. Letermovir Prophylaxis for Cytomegalovirus in Hematopoietic-Cell Transplantation. *New England Journal of Medicine,* 377**,** 2433-2444.

MCGEOCH, D. J., COOK, S., DOLAN, A., JAMIESON, F. E. & TELFORD, E. A. 1995. Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *J Mol Biol,* 247**,** 443-58.

MCLAUGHLIN-TAYLOR, E., PANDE, H., FORMAN, S. J., TANAMACHI, B., LI, C. R., ZAIA, J. A., GREENBERG, P. D. & RIDDELL, S. R. 1994. Identification of the major late human cytomegalovirus matrix protein pp65 as a target antigen for CD8+ virus-specific cytotoxic T lymphocytes. *J Med Virol,* 43**,** 103-10.

MCSHARRY, B. P., JONES, C. J., SKINNER, J. W., KIPLING, D. & WILKINSON, G. W. 2001. Human telomerase reverse transcriptase-immortalized MRC-5 and HCA2 human fibroblasts are fully permissive for human cytomegalovirus. *J Gen Virol,* 82**,** 855-63.

MEGJUGORAC, N. J., YOUNG, H. A., AMRUTE, S. B., OLSHALSKY, S. L. & FITZGERALD-BOCARSLY, P. 2004. Virally stimulated plasmacytoid dendritic cells produce chemokines and induce migration of T and NK cells. *J Leukoc Biol,* 75**,** 504-14.

MEYER, H. H., RIPALTI, A., LANDINI, M. P., RADSAK, K., KERN, H. F. & HENSEL, G. M. 1997. Human cytomegalovirus late-phase maturation is blocked by stably expressed UL32 antisense mRNA in astrocytoma cells. *J Gen Virol,* 78.

MILBRADT, J., AUEROCHS, S. & MARSCHALL, M. 2007. Cytomegaloviral proteins pUL50 and pUL53 are associated with the nuclear lamina and interact with cellular protein kinase C. *J Gen Virol,* 88**,** 2642-50.

MOCARSKI, E., SHENK, T., GRIFFITHS, P. D. & PASS, R. F. 2013. Cytomegaloviruses. *In:* FIELDS, B. N., KNIPE, D. M. & HOWLEY, P. M. (eds.) *Fields Virology.* 6th ed. Philadelphia: Lippincott Williams & Wilkins.

MOCARSKI, E. S. & COURCELLE, C. T. 2001. Cytomegaloviruses and their replication. *In:* D. M. KNIPE, P. M. H., D. E. GRIFFIN, R. A. LAMB, M. A. MARTIN, B. ROIZMAN, AND S. E. STRAUS (ed.) *Fields virology.* 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins.

MODLIN, J. F., ARVIN, A. M., FAST, P., MYERS, M., PLOTKIN, S. & RABINOVICH, R. 2004. Vaccine Development to Prevent Cytomegalovirus Disease: Report from the National Vaccine Advisory Committee. *Clinical Infectious Diseases,* 39**,** 233-239.

MOORE, P. S., GAO, S. J., DOMINGUEZ, G., CESARMAN, E., LUNGU, O., KNOWLES, D. M., GARBER, R., PELLETT, P. E., MCGEOCH, D. J. & CHANG, Y. 1996. Primary characterization of a herpesvirus agent associated with Kaposi's sarcomae. *J Virol,* 70**,** 549-58.

MURALIDHARAN, S. & MANDREKAR, P. 2013. Cellular stress response and innate immune signaling: integrating pathways in host defense and inflammation. *Journal of Leukocyte Biology,* 94**,** 1167-1184.

MURPHY, E., RIGOUTSOS, I., SHIBUYA, T. & SHENK, T. E. 2003. Reevaluation of human cytomegalovirus coding potential. *Proc Natl Acad Sci U S A,* 100.

MURPHY, E. & SHENK, T. 2008. Human cytomegalovirus genome. *In:* MURPHY, E. & SHENK, T. (eds.) *Human Cytomegaloviruses (Current Topics in Microbiology and Immunology).* Berlin Heidelberg: Springer-Verlag.

MURRELL, I. 2014. *Developing a Human Cytomegalovirus strain for better in vitro research.* PhD thesis, Cardiff University.

MURRELL, I., BEDFORD, C., LADELL, K., MINERS, K. L., PRICE, D. A., TOMASEC, P., WILKINSON, G. W. G. & STANTON, R. J. 2017. The pentameric complex drives immunologically covert cell-cell transmission of wild-type human cytomegalovirus. *Proc Natl Acad Sci U S A,* 114**,** 6104-6109.

MURRELL, I., TOMASEC, P., WILKIE, G. S., DARGAN, D. J., DAVISON, A. J. & STANTON, R. J. 2013. Impact of sequence variation in the UL128 locus on production of human cytomegalovirus in fibroblast and epithelial cells. *J Virol,* 87**,** 10489-500.

MYERSON, D., HACKMAN, R. C., NELSON, J. A., WARD, D. C. & MCDOUGALL, J. K. 1984. Widespread presence of histologically occult cytomegalovirus. *Hum Pathol,* 15**,** 430-9.

NAKAI, K. & HORTON, P. 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci,* 24**,** 34-6.

NAKAMAE, H., STORER, B., SANDMAIER, B. M., MALONEY, D. G., DAVIS, C., COREY, L., STORB, R. & BOECKH, M. 2011. Cytopenias after day 28 in allogeneic hematopoietic cell transplantation: impact of recipient/donor factors, transplant conditions and myelotoxic drugs. *Haematologica,* 96**,** 1838-1845.

NAKAMURA, R., LA ROSA, C., LONGMATE, J., DRAKE, J., SLAPE, C., ZHOU, Q., LAMPA, M. G., O'DONNELL, M., CAI, J.-L., FAROL, L., SALHOTRA, A., SNYDER, D. S., ALDOSS, I., FORMAN, S. J., MILLER, J. S., ZAIA, J. A. & DIAMOND, D. J. 2016. Cytomegalovirus chimeric epitope vaccine supplemented with PF03512676 (CMVPepVax) in allogeneic hematopoietic stem cell transplantation: viremia, immunogenicity and

survival outcomes in a randomised phase 1b trial. *The Lancet. Haematology,* 3**,** e87-e98.

NATHAN, C. & SHILOH, M. U. 2000. Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc Natl Acad Sci U S A,* 97**,** 8841-8.

NELMS, B. L. & LABOSKY, P. A. 2011. A predicted hairpin cluster correlates with barriers to PCR, sequencing and possibly BAC recombineering. *Scientific Reports,* 1**,** 106.

NGUYEN, K. B., SALAZAR-MATHER, T. P., DALOD, M. Y., VAN DEUSEN, J. B., WEI, X. Q., LIEW, F. Y., CALIGIURI, M. A., DURBIN, J. E. & BIRON, C. A. 2002. Coordinated and distinct roles for IFN-alpha beta, IL-12, and IL-15 regulation of NK cell responses to viral infection. *J Immunol,* 169**,** 4279-87.

NIGRO, G., ADLER, S. P., LA TORRE, R. & BEST, A. M. 2005. Passive immunization during pregnancy for congenital cytomegalovirus infection. *N Engl J Med,* 353**,** 1350-62.

NORIEGA, V., REDMANN, V., GARDNER, T. & TORTORELLA, D. 2012a. Diverse immune evasion strategies by human cytomegalovirus. *Immunol Res,* 54**,** 140-51.

NORIEGA, V. M., HESSE, J., GARDNER, T. J., BESOLD, K., PLACHTER, B. & TORTORELLA, D. 2012b. Human Cytomegalovirus US3 Modulates Destruction of MHC Class I Molecules. *Molecular Immunology,* 51**,** 245-253.

NOYOLA, D. E., DEMMLER, G. J., WILLIAMSON, W. D., GRIESSER, C., SELLERS, S., LLORENTE, A., LITTMAN, T., WILLIAMS, S., JARRETT, L. & YOW, M. D. 2000. Cytomegalovirus urinary excretion and long term outcome in children with congenital cytomegalovirus infection. Congenital CMV Longitudinal Study Group. *Pediatr Infect Dis J,* 19**,** 505-10.

OGAWA-GOTO, K., TANAKA, K., GIBSON, W., MORIISHI, E., MIURA, Y., KURATA, T., IRIE, S. & SATA, T. 2003. Microtubule Network Facilitates Nuclear Targeting of Human Cytomegalovirus Capsid. *Journal of Virology,* 77**,** 8541-8547.

OGAWA, H., SUZUTAN, T., BABA, Y., KOYANO, S., NOZAWA, N., ISHIBASHI, K., FUJIEDA, K., INOUE, N. & OMORI, K. 2007. Etiology of Severe Sensorineural Hearing Loss in Children: Independent Impact of Congenital Cytomegalovirus Infection and GJB2 Mutations. *The Journal of Infectious Diseases,* 195**,** 782-788.

OMASITS, U., AHRENS, C. H., MÜLLER, S. & WOLLSCHEID, B. 2014. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics,* 30**,** 884-886.

ORANGE, J. S. 2013. Natural killer cell deficiency. *The Journal of allergy and clinical immunology,* 132**,** 515-526.

PANDE, N. T., POWERS, C., AHN, K. & FRÜH, K. 2005. Rhesus Cytomegalovirus Contains Functional Homologues of US2, US3, US6, and US11. *Journal of Virology,* 79**,** 5786-5798.

PARI, G. S. 2008. Nuts and bolts of human cytomegalovirus lytic DNA replication. *Curr Top Microbiol Immunol,* 325**,** 153-66.

PARI, G. S. & ANDERS, D. G. 1993. Eleven loci encoding trans-acting factors are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA replication. *Journal of Virology,* 67**,** 6979-6988.

PARK, B., SPOONER, E., HOUSER, B. L., STROMINGER, J. L. & PLOEGH, H. L. 2010. The HCMV membrane glycoprotein US10 selectively targets HLA-G for degradation. *J Exp Med,* 207**,** 2033-41.

PASS, R. F., DULIEGE, A. M., BOPPANA, S., SEKULOVICH, R., PERCELL, S., BRITT, W. & BURKE, R. L. 1999. A subunit cytomegalovirus vaccine based on recombinant envelope glycoprotein B and a new adjuvant. *J Infect Dis,* 180**,** 970-5.

PASS, R. F., FOWLER, K. B., BOPPANA, S. B., BRITT, W. J. & STAGNO, S. 2006. Congenital cytomegalovirus infection following first trimester maternal infection: symptoms at birth and outcome. *J Clin Virol,* 35**,** 216-20.

PASS, R. F., STAGNO, S., MYERS, G. J. & ALFORD, C. A. 1980. Outcome of symptomatic congenital cytomegalovirus infection: results of long-term longitudinal follow-up. *Pediatrics,* 66**,** 758-62.

PASS, R. F., ZHANG, C., EVANS, A., SIMPSON, T., ANDREWS, W., HUANG, M.-L., COREY, L., HILL, J., DAVIS, E., FLANIGAN, C. & CLOUD, G. 2009. Vaccine Prevention of Maternal Cytomegalovirus Infection. *The New England journal of medicine,* 360**,** 1191-1199.

PEGRAM, H. J., ANDREWS, D. M., SMYTH, M. J., DARCY, P. K. & KERSHAW, M. H. 2011. Activating and inhibitory receptors of natural killer cells. *Immunol Cell Biol,* 89**,** 216-224.

PELLETT, P. E. & ROIZMAN, B. 2013. Herpesviridae. *In:* FIELDS, B. N., KNIPE, D. M. & HOWLEY, P. M. (eds.) *Fields Virology.* 6th ed. Philadelphia: Lippincott Williams & Wilkins.

PENFOLD, M. E., DAIRAGHI, D. J., DUKE, G. M., SAEDERUP, N., MOCARSKI, E. S., KEMBLE, G. W. & SCHALL, T. J. 1999. Cytomegalovirus encodes a potent alpha chemokine. *Proc Natl Acad Sci U S A,* 96**,** 9839-44.

PETER, M. E. & KRAMMER, P. H. 2003. The CD95(APO-1/Fas) DISC and beyond. *Cell Death Differ,* 10**,** 26-35.

PFEFFER, K. 2003. Biological functions of tumor necrosis factor cytokines and their receptors. *Cytokine Growth Factor Rev,* 14**,** 185-91.

PLAFKER, S. M. & GIBSON, W. 1998. Cytomegalovirus assembly protein precursor and proteinase precursor contain two nuclear localization signals that mediate their own nuclear translocation and that of the major capsid protein. *J Virol,* 72**,** 7722-32.

PLOTKIN, S. A., FURUKAWA, T., ZYGRAICH, N. & HUYGELEN, C. 1975. Candidate cytomegalovirus strain for human vaccination. *Infection and Immunity,* 12**,** 521-527.

PLOTKIN, S. A., HIGGINS, R., KURTZ, J. B., MORRIS, P. J., CAMPBELL, D. A., JR., SHOPE, T. C., SPECTOR, S. A. & DANKNER, W. M. 1994. Multicenter trial of Towne strain attenuated virus vaccine in seronegative renal transplant recipients. *Transplantation,* 58**,** 1176-8.

PLOTKIN, S. A., SMILEY, M. L., FRIEDMAN, H. M., STARR, S. E., FLEISHER, G. R., WLODAVER, C., DAFOE, D. C., FRIEDMAN, A. D., GROSSMAN, R. A. & BARKER, C. F. 1984. Towne-vaccine-induced prevention of cytomegalovirus disease after renal transplants. *Lancet,* 1**,** 528-30.

PLOTKIN, S. A., STARR, S. E., FRIEDMAN, H. M., BRAYMAN, K., HARRIS, S., JACKSON, S., TUSTIN, N. B., GROSSMAN, R., DAFOE, D. & BARKER, C. 1991. Effect of Towne live virus vaccine on cytomegalovirus disease after renal transplant. A controlled trial. *Ann Intern Med,* 114**,** 525-31.

POOLE, E., AVDIC, S., HODKINSON, J., JACKSON, S., WILLS, M., SLOBEDMAN, B. & SINCLAIR, J. 2014. Latency-associated viral interleukin-10 (IL-10) encoded by human cytomegalovirus modulates cellular IL-10 and CCL8 Secretion during latent infection through changes in the cellular microRNA hsa-miR-92a. *J Virol,* 88**,** 13947-55.

POOLE, E., KING, C. A., SINCLAIR, J. H. & ALCAMI, A. 2006. The UL144 gene product of human cytomegalovirus activates NFκB via a TRAF6-dependent mechanism. *The EMBO Journal,* 25**,** 4390-4399.

PRINCE, V. E. & PICKETT, F. B. 2002. Splitting pairs: the diverging fates of duplicated genes. *Nat Rev Genet,* 3**,** 827-37.

PROD'HOMME, V., SUGRUE, D. M., STANTON, R. J., NOMOTO, A., DAVIES, J., RICKARDS, C. R., COCHRANE, D., MOORE, M., WILKINSON, G. W. G. & TOMASEC, P. 2010. Human cytomegalovirus UL141 promotes efficient downregulation of the natural killer cell activating ligand CD112. *The Journal of General Virology,* 91**,** 2034-2039.

PROD'HOMME, V., GRIFFIN, C., AICHELER, R. J., WANG, E. C. Y., MCSHARRY, B. P., RICKARDS, C. R., STANTON, R. J., BORYSIEWICZ, L. K., LÓPEZ-BOTET, M., WILKINSON, G. W. G. & TOMASEC, P. 2007. The Human Cytomegalovirus MHC Class I Homolog UL18 Inhibits LIR-1(+) but Activates LIR-1(−) NK Cells. *Journal of immunology (Baltimore, Md. : 1950),* 178**,** 4473-4481.

PROD'HOMME, V., TOMASEC, P., CUNNINGHAM, C., LEMBERG, M. K., STANTON, R. J., MCSHARRY, B. P., WANG, E. C. Y., CUFF, S., MARTOGLIO, B., DAVISON, A. J., BRAUD, V. M. & WILKINSON, G. W. G. 2012. Human Cytomegalovirus UL40 Signal Peptide Regulates Cell Surface Expression of the Natural Killer Cell Ligands HLA-E and gpUL18. *Journal of immunology (Baltimore, Md. : 1950),* 188**,** 2794-2804.

QUINNAN, G. V., JR., DELERY, M., ROOK, A. H., FREDERICK, W. R., EPSTEIN, J. S., MANISCHEWITZ, J. F., JACKSON, L., RAMSEY, K. M., MITTAL, K., PLOTKIN, S. A. & ET AL. 1984. Comparative virulence and immunogenicity of the Towne strain and a nonattenuated strain of cytomegalovirus. *Ann Intern Med,* 101**,** 478-83.

RABELLINO, A. & SCAGLIONI, P. P. 2013. PML Degradation: Multiple Ways to Eliminate PML. *Frontiers in Oncology,* 3.

RATH, A., GLIBOWICKA, M., NADEAU, V. G., CHEN, G. & DEBER, C. M. 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences,* 106**,** 1760-1765.

REEVES, M., MURPHY, J., GREAVES, R., FAIRLEY, J., BREHM, A. & SINCLAIR, J. 2006. Autorepression of the human cytomegalovirus major immediate-early promoter/enhancer at late times of infection is mediated by the recruitment of chromatin remodeling enzymes by IE86. *J Virol,* 80**,** 9998-10009.

REEVES, M. B., LEHNER, P. J., SISSONS, J. G. & SINCLAIR, J. H. 2005a. An in vitro model for the regulation of human cytomegalovirus latency and reactivation in dendritic cells by chromatin remodelling. *J Gen Virol,* 86**,** 2949-54.

REEVES, M. B., MACARY, P. A., LEHNER, P. J., SISSONS, J. G. P. & SINCLAIR, J. H. 2005b. Latency, chromatin remodeling, and reactivation of human cytomegalovirus in the dendritic cells of healthy carriers. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 4140-4145.

RIBBERT, D. 1904. Uber protozoenartige zellen in der niere eines syphilitischen neugoborenen und in der parotis von kindern. *Zentralbl. Allg. Pathol.***,** 945-948.

RIGOUTSOS, I., NOVOTNY, J., HUYNH, T., CHIN-BOW, S. T., PARIDA, L., PLATT, D., COLEMAN, D. & SHENK, T. 2003. In silico pattern-based analysis of the human cytomegalovirus genome. *J Virol,* 77**,** 4326-44.

ROIZMAN, B. & PELLETT, P. E. 2001. The Family Herpesviridae: A brief introduction. . *In:* KNIPE, D. M., GRIFFIN, D. E., LAMB, R. A., MARTIN, M. A., ROIZMAN, B. & STRAUS, S. E. (eds.) *Fields virology.* 4 ed. Philadelphia: Lippincott Williams & Wilkins.

ROJAS-RIVERA, D., ARMISEN, R., COLOMBO, A., MARTINEZ, G., EGUIGUREN, A. L., DIAZ, A., KIVILUOTO, S., RODRIGUEZ, D., PATRON, M., RIZZUTO, R., BULTYNCK, G., CONCHA, M. L., SIERRALTA, J., STUTZIN, A. & HETZ, C. 2012. TMBIM3/GRINA is a novel unfolded protein response (UPR) target gene that controls apoptosis through the modulation of ER calcium homeostasis. *Cell Death Differ,* 19**,** 1013-26.

ROJAS-RIVERA, D. & HETZ, C. 2015. TMBIM protein family: ancestral regulators of cell death. *Oncogene,* 34**,** 269-80.

ROLLE, A., MOUSAVI-JAZI, M., ERIKSSON, M., ODEBERG, J., SODERBERG-NAUCLER, C., COSMAN, D., KARRE, K. & CERBONI, C. 2003. Effects of human cytomegalovirus

infection on ligands for the activating NKG2D receptor of NK cells: up-regulation of UL16-binding protein (ULBP)1 and ULBP2 is counteracted by the viral UL16 protein. *J Immunol,* 171**,** 902-8.

ROSENKILDE, M. M., WALDHOER, M., LUTTICHAU, H. R. & SCHWARTZ, T. W. 2001. Virally encoded 7TM receptors. *Oncogene,* 20**,** 1582-93.

ROSS, S. A., FOWLER, K. B., ASHRITH, G., STAGNO, S., BRITT, W. J., PASS, R. F. & BOPPANA, S. B. 2006. Hearing loss in children with congenital cytomegalovirus infection born to mothers with preexisting immunity. *The Journal of Pediatrics,* 148**,** 332-336.

ROSSINI, G., CERBONI, C., SANTONI, A., LANDINI, M. P., LANDOLFO, S., GATTI, D., GRIBAUDO, G. & VARANI, S. 2012. Interplay between Human Cytomegalovirus and Intrinsic/Innate Host Responses: A Complex Bidirectional Relationship. *Mediators of Inflammation,* 2012**,** 607276.

ROWE, W. P., HUEBNER, R. J., GILMORE, L. K., PARROTT, R. H. & WARD, T. G. 1953. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc Soc Exp Biol Med,* 84**,** 570-3.

RUDD, P. M., ELLIOTT, T., CRESSWELL, P., WILSON, I. A. & DWEK, R. A. 2001. Glycosylation and the immune system. *Science,* 291**,** 2370-6.

RYCKMAN, B. J., JARVIS, M. A., DRUMMOND, D. D., NELSON, J. A. & JOHNSON, D. C. 2006. Human Cytomegalovirus Entry into Epithelial and Endothelial Cells Depends on Genes UL128 to UL150 and Occurs by Endocytosis and Low-pH Fusion. *Journal of Virology,* 80**,** 710-722.

SAFFERT, R. T. & KALEJTA, R. F. 2006. Inactivating a Cellular Intrinsic Immune Defense Mediated by Daxx Is the Mechanism through Which the Human Cytomegalovirus pp71 Protein Stimulates Viral Immediate-Early Gene Expression. *Journal of Virology,* 80**,** 3863-3871.

SALSMAN, J., JAGANNATHAN, M., PALADINO, P., CHAN, P.-K., DELLAIRE, G., RAUGHT, B. & FRAPPIER, L. 2012. Proteomic Profiling of the Human Cytomegalovirus UL35 Gene Products Reveals a Role for UL35 in the DNA Repair Response. *Journal of Virology,* 86**,** 806-820.

SALSMAN, J., WANG, X. & FRAPPIER, L. 2011. Nuclear body formation and PML body remodeling by the human cytomegalovirus protein UL35. *Virology,* 414**,** 119-129.

SALSMAN, J., ZIMMERMAN, N., CHEN, T., DOMAGALA, M. & FRAPPIER, L. 2008. Genome-Wide Screen of Three Herpesviruses for Protein Subcellular Localization and Alteration of PML Nuclear Bodies. *PLoS Pathogens,* 4**,** e1000100.

SALVANT, B. S., FORTUNATO, E. A. & SPECTOR, D. H. 1998. Cell Cycle Dysregulation by Human Cytomegalovirus: Influence of the Cell Cycle Phase at the Time of Infection and Effects on Cyclin Transcription. *Journal of Virology,* 72**,** 3729-3741.

SANCHEZ, V., ANGELETTI, P. C., ENGLER, J. A. & BRITT, W. J. 1998. Localization of Human Cytomegalovirus Structural Proteins to the Nuclear Matrix of Infected Human Fibroblasts. *Journal of Virology,* 72**,** 3321-3329.

SANCHEZ, V., GREIS, K. D., SZTUL, E. & BRITT, W. J. 2000a. Accumulation of Virion Tegument and Envelope Proteins in a Stable Cytoplasmic Compartment during Human Cytomegalovirus Replication: Characterization of a Potential Site of Virus Assembly *J Virol,* 74**,** 975-986.

SANCHEZ, V., SZTUL, E. & BRITT, W. J. 2000b. Human cytomegalovirus pp28 (UL99) localizes to a cytoplasmic compartment which overlaps the endoplasmic reticulum-golgi-intermediate compartment. *J Virol,* 74**,** 3842-51.

SARISKY, R. T. & HAYWARD, G. S. 1996. Evidence that the UL84 gene product of human cytomegalovirus is essential for promoting oriLyt-dependent DNA replication and formation of replication compartments in cotransfection assays. *J Virol,* 70**,** 7398-413.

SCARBOROUGH, J. A., PAUL, J. R. & SPENCER, J. V. 2017. Evolution of the ability to modulate host chemokine networks via gene duplication in human cytomegalovirus (HCMV). *Infect Genet Evol,* 51**,** 46-53.

SCHAUFLINGER, M., VILLINGER, C., MERTENS, T., WALTHER, P. & VON EINEM, J. 2013. Analysis of human cytomegalovirus secondary envelopment by advanced electron microscopy. *Cell Microbiol,* 15**,** 305-14.

SCHNELL, J. D. & HICKE, L. 2003. Non-traditional Functions of Ubiquitin and Ubiquitin-binding Proteins. *Journal of Biological Chemistry,* 278**,** 35857-35860.

SCRIVANO, L., SINZGER, C., NITSCHKO, H., KOSZINOWSKI, U. H. & ADLER, B. 2011. HCMV spread and cell tropism are determined by distinct virus populations. *PLoS Pathog,* 7**,** e1001256.

SEGLEN, P. O., GRINDE, B. & SOLHEIM, A. E. 1979. Inhibition of the lysosomal pathway of protein degradation in isolated rat hepatocytes by ammonia, methylamine, chloroquine and leupeptin. *Eur J Biochem,* 95**,** 215-25.

SEIDEL, E., LE, V. T., BAR-ON, Y., TSUKERMAN, P., ENK, J., YAMIN, R., STEIN, N., SCHMIEDEL, D., OIKNINE DJIAN, E., WEISBLUM, Y., TIROSH, B., STASTNY, P., WOLF, D. G., HENGEL, H. & MANDELBOIM, O. 2015. Dynamic Co-evolution of Host and Pathogen: HCMV Downregulates the Prevalent Allele MICA *008 to Escape Elimination by NK Cells. *Cell Rep*.

SEIRAFIAN, S. 2012. *An Analysis of Human Cytomegalovirus Gene Usage.* PhD thesis, Cardiff University.

SEIRAFIAN, S., PROD'HOMME, V., SUGRUE, D., DAVIES, J., FIELDING, C., TOMASEC, P. & WILKINSON, G. W. 2014. Human cytomegalovirus suppresses Fas expression and function. *J Gen Virol,* 95**,** 933-9.

SEKULIN, K., GORZER, I., HEISS-CZEDIK, D. & PUCHHAMMER-STOCKL, E. 2007. Analysis of the variability of CMV strains in the RL11D domain of the RL11 multigene family. *Virus Genes,* 35**,** 577-83.

SEO, S. & BOECKH, M. 2013. Clinical Cytomegalovirus Research: Haemopoietic Cell Transplant. *In:* REDDEHASE, M. J. (ed.) *Cytomegaloviruses: from Molecular Pathogenesis to Intervention* Norfolk, UK: Caister Academic Press.

SEYFRIED, N. T., XU, P., DUONG, D. M., CHENG, D., HANFELT, J. & PENG, J. 2008. Systematic approach for validating the ubiquitinated proteome. *Anal Chem,* 80**,** 4161-9.

SHAO, S. & HEGDE, R. S. 2011. Membrane Protein Insertion at the Endoplasmic Reticulum. *Annual review of cell and developmental biology,* 27**,** 25-56.

SHARMA, M., KAMIL, J. P., COUGHLIN, M., REIM, N. I. & COEN, D. M. 2014. Human Cytomegalovirus UL50 and UL53 Recruit Viral Protein Kinase UL97, Not Protein Kinase C, for Disruption of Nuclear Lamina and Nuclear Egress in Infected Cells. *Journal of Virology,* 88**,** 249-262.

SHIKHAGAIE, M., MERCÉ-MALDONADO, E., ISERN, E., MUNTASELL, A., ALBÀ, M. M., LÓPEZ-BOTET, M., HENGEL, H. & ANGULO, A. 2012. The human cytomegalovirus-specific UL1 gene encodes a late-phase glycoprotein incorporated in the virion envelope. *Journal of virology,* 86**,** 4091-4101.

SIJMONS, S., THYS, K., MBONG NGWESE, M., VAN DAMME, E., DVORAK, J., VAN LOOCK, M., LI, G., TACHEZY, R., BUSSON, L., AERSSENS, J., VAN RANST, M. & MAES, P. 2015. High-throughput analysis of human cytomegalovirus genome diversity highlights the widespread occurrence of gene-disrupting mutations and pervasive recombination. *J Virol*.

SIJMONS, S., VAN RANST, M. & MAES, P. 2014. Genomic and Functional Characteristics of Human Cytomegalovirus Revealed by Next-Generation Sequencing. *Viruses,* 6**,** 1049-1072.

SIMMEN, K. A., SINGH, J., LUUKKONEN, B. G. M., LOPPER, M., BITTNER, A., MILLER, N. E., JACKSON, M. R., COMPTON, T. & FRÜH, K. 2001. Global modulation of cellular transcription by human cytomegalovirus is initiated by viral glycoprotein B. *Proceedings of the National Academy of Sciences,* 98**,** 7140-7145.

SINCLAIR, J. 2008. Human cytomegalovirus: Latency and reactivation in the myeloid lineage. *Journal of Clinical Virology,* 41**,** 180-185.

SINCLAIR, J. & SISSONS, P. 2006. Latency and reactivation of human cytomegalovirus. *J Gen Virol,* 87**,** 1763-79.

SINZGER, C., GREFTE, A., PLACHTER, B., GOUW, A. S., THE, T. H. & JAHN, G. 1995. Fibroblasts, epithelial cells, endothelial cells and smooth muscle cells are major targets of human cytomegalovirus infection in lung and gastrointestinal tissues. *J Gen Virol,* 76 (Pt 4)**,** 741-50.

SINZGER, C., HAHN, G., DIGEL, M., KATONA, R., SAMPAIO, K. L., MESSERLE, M., HENGEL, H., KOSZINOWSKI, U., BRUNE, W. & ADLER, B. 2008. Cloning and sequencing of a highly productive, endotheliotropic virus strain derived from human cytomegalovirus TB40/E. *J Gen Virol,* 89**,** 359-68.

SKALETSKAYA, A., BARTLE, L. M., CHITTENDEN, T., MCCORMICK, A. L., MOCARSKI, E. S. & GOLDMACHER, V. S. 2001. A cytomegalovirus-encoded inhibitor of apoptosis that suppresses caspase-8 activation. *Proc Natl Acad Sci U S A,* 98**,** 7829-34.

SLAVULJICA, I., KRMPOTIĆ, A. & JONJIĆ, S. 2011. Manipulation of NKG2D Ligands by Cytomegaloviruses: Impact on Innate and Adaptive Immune Response. *Frontiers in Immunology,* 2**,** 85.

SLOBEDMAN, B. & MOCARSKI, E. S. 1999. Quantitative Analysis of Latent Human Cytomegalovirus. *Journal of Virology,* 73**,** 4806-4812.

SMITH, I. L., TASKINTUNA, I., RAHHAL, F. M., POWELL, H. C., AI, E., MUELLER, A. J., SPECTOR, S. A. & FREEMAN, W. R. 1998. Clinical Failure of CMV Retinitis With Intravitreal Cidofovir Is Associated With Antiviral Resistance. *Arch. Ophthalmol,* 116.

SMITH, J. A., TURNER, M. J., DELAY, M. L., KLENK, E. I., SOWDERS, D. P. & COLBERT, R. A. 2008. Endoplasmic reticulum stress and the unfolded protein response are linked to synergistic IFN-beta induction via X-box binding protein 1. *Eur J Immunol,* 38**,** 1194-203.

SMITH, L. R., WLOCH, M. K., CHAPLIN, J. A., GERBER, M. & ROLLAND, A. P. 2013a. Clinical Development of a Cytomegalovirus DNA Vaccine: From Product Concept to Pivotal Phase 3 Trial. *Vaccines (Basel),* 1**,** 398-414.

SMITH, W., TOMASEC, P., AICHELER, R., LOEWENDORF, A., NEMČOVIČOVÁ, I., WANG, EDDIE C., STANTON, RICHARD J., MACAULEY, M., NORRIS, P., WILLEN, L., RUCKOVA, E., NOMOTO, A., SCHNEIDER, P., HAHN, G., ZAJONC, DIRK M., WARE, CARL F., WILKINSON, GAVIN W. & BENEDICT, CHRIS A. 2013b. Human Cytomegalovirus Glycoprotein UL141 Targets the TRAIL Death Receptors to Thwart Host Innate Antiviral Defenses. *Cell Host & Microbe,* 13**,** 324-335.

SMYTH, M. J., KELLY, J. M., SUTTON, V. R., DAVIS, J. E., BROWNE, K. A., SAYERS, T. J. & TRAPANI, J. A. 2001. Unlocking the secrets of cytotoxic granule proteins. *J Leukoc Biol,* 70**,** 18-29.

SÖDERBERG-NAUCLÉR, C. & NELSON, J. A. 1999. Human Cytomegalovirus Latency and Reactivation – A Delicate Balance between the Virus and Its Host's Immune System. *Intervirology***,** 314-321.

SOROCEANU, L., AKHAVAN, A. & COBBS, C. S. 2008. Platelet-derived growth factor-alpha receptor activation is required for human cytomegalovirus infection. *Nature,* 455**,** 391-5.

SPAETE, R. R., GEHRZ, R. C. & LANDINI, M. P. 1994. Human cytomegalovirus structural proteins. *J Gen Virol,* 75 (Pt 12)**,** 3287-308.

SPEAR, P. G. & LONGNECKER, R. 2003. Herpesvirus entry: an update. *J Virol,* 77.

SPECTOR, D. J. & YETMING, K. 2010. UL84-independent replication of human cytomegalovirus strain TB40/E. *Virology,* 407**,** 171-177.

SPECTOR, S. A., WONG, R., HSIA, K., PILCHER, M. & STEMPIEN, M. J. 1998. Plasma cytomegalovirus (CMV) DNA load predicts CMV disease and survival in AIDS patients. *Journal of Clinical Investigation,* 101**,** 497-502.

SPENCER, J. V., LOCKRIDGE, K. M., BARRY, P. A., LIN, G., TSANG, M., PENFOLD, M. E. & SCHALL, T. J. 2002. Potent immunosuppressive activities of cytomegalovirus-encoded interleukin-10. *J Virol,* 76**,** 1285-92.

SPITS, H., LANIER, L. & PHILLIPS, J. 1995. Development of human T and natural killer cells. *Blood,* 85**,** 2654-2670.

STAGNO, S., PASS, R. F., CLOUD, G. & ET AL. 1986. Primary cytomegalovirus infection in pregnancy: Incidence, transmission to fetus, and clinical outcome. *JAMA,* 256**,** 1904-1908.

STANTON, R., WESTMORELAND, D., FOX, J. D., DAVISON, A. J. & WILKINSON, G. W. 2005. Stability of human cytomegalovirus genotypes in persistently infected renal transplant recipients. *J Med Virol,* 75**,** 42-6.

STANTON, R. J., BALUCHOVA, K., DARGAN, D. J., CUNNINGHAM, C., SHEEHY, O., SEIRAFIAN, S., MCSHARRY, B. P., NEALE, M. L., DAVIES, J. A. & TOMASEC, P. 2010. Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication. *The Journal of clinical investigation,* 120**,** 3191.

STANTON, R. J., MCSHARRY, B. P., ARMSTRONG, M., TOMASEC, P. & WILKINSON, G. W. G. 2008. Re-engineering adenovirus vector systems to enable high-throughput analyses of gene function. *Biotechniques. The International Journal of Life Science Methods,* 45**,** 659-668.

STANTON, R. J., PROD'HOMME, V., PURBHOO, M. A., MOORE, M., AICHELER, R. J., HEINZMANN, M., BAILER, S. M., HAAS, J., ANTROBUS, R., WEEKES, M. P., LEHNER, P. J., VOJTESEK, B., MINERS, K. L., MAN, S., WILKIE, G. S., DAVISON, A. J., WANG, E. C., TOMASEC, P. & WILKINSON, G. W. 2014. HCMV pUL135 remodels the actin cytoskeleton to impair immune recognition of infected cells. *Cell Host Microbe,* 16**,** 201-14.

STARAS, S. A., DOLLARD, S. C., RADFORD, K. W., FLANDERS, W. D., PASS, R. F. & CANNON, M. J. 2006. Seroprevalence of cytomegalovirus infection in the United States, 1988-1994. *Clin Infect Dis,* 43**,** 1143-51.

STAUDT, C., PUISSANT, E. & BOONEN, M. 2017. Subcellular Trafficking of Mammalian Lysosomal Proteins: An Extended View. *International Journal of Molecular Sciences,* 18**,** 47.

STENBERG, R. M., THOMSEN, D. R. & STINSKI, M. F. 1984. Structural analysis of the major immediate early gene of human cytomegalovirus. *Journal of Virology,* 49**,** 190-199.

STERN-GINOSSAR, N., ELEFANT, N., ZIMMERMANN, A., WOLF, D. G., SALEH, N., BITON, M., HORWITZ, E., PROKOCIMER, Z., PRICHARD, M., HAHN, G., GOLDMAN-WOHL, D., GREENFIELD, C., YAGEL, S., HENGEL, H., ALTUVIA, Y., MARGALIT, H. & MANDELBOIM, O. 2007. Host immune system gene targeting by a viral miRNA. *Science,* 317**,** 376-81.

STERN-GINOSSAR, N., WEISBURD, B., MICHALSKI, A., LE, V. T., HEIN, M. Y., HUANG, S. X., MA, M., SHEN, B., QIAN, S. B. & HENGEL, H. 2012. Decoding human cytomegalovirus. *Science,* 338.

STINSKI, M. F. 1977. Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. *J Virol,* 23**,** 751-67.

STINSKI, M. F., THOMSEN, D. R., STENBERG, R. M. & GOLDSTEIN, L. C. 1983. Organization and expression of the immediate early genes of human cytomegalovirus. *Journal of Virology,* 46**,** 1-14.

STRANG, B. L. 2015. Viral and cellular subnuclear structures in human cytomegalovirus-infected cells. *Journal of General Virology,* 96**,** 239-253.

SUNG, C. K., LI, H., CLAVERYS, J. P. & MORRISON, D. A. 2001. An rpsL Cassette, Janus, for Gene Replacement through Negative Selection in Streptococcus pneumoniae. *Applied and Environmental Microbiology,* 67**,** 5190-5196.

SUTHERLAND, C. L., CHALUPNY, N. J., SCHOOLEY, K., VANDENBOS, T., KUBIN, M. & COSMAN, D. 2002. UL16-binding proteins, novel MHC class I-related proteins, bind to NKG2D and activate multiple signaling pathways in primary NK cells. *J Immunol,* 168**,** 671-9.

SYLWESTER, A. W., MITCHELL, B. L., EDGAR, J. B., TAORMINA, C., PELTE, C., RUCHTI, F., SLEATH, P. R., GRABSTEIN, K. H., HOSKEN, N. A., KERN, F., NELSON, J. A. & PICKER, L. J. 2005. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. *J Exp Med,* 202**,** 673-85.

TADAGAKI, K., TUDOR, D., GBAHOU, F., TSCHISCHE, P., WALDHOER, M., BOMSEL, M., JOCKERS, R. & KAMAL, M. 2012. Human cytomegalovirus-encoded UL33 and UL78 heteromerize with host CCR5 and CXCR4 impairing their HIV coreceptor activity. *Blood,* 119**,** 4908-18.

TANDON, R. & MOCARSKI, E. S. 2012. Viral and host control of cytomegalovirus maturation. *Trends in microbiology,* 20**,** 392-401.

TANDON, R., MOCARSKI, E. S. & CONWAY, J. F. 2015. The A, B, Cs of Herpesvirus Capsids. *Viruses,* 7**,** 899-914.

TAVALAI, N., PAPIOR, P., RECHTER, S., LEIS, M. & STAMMINGER, T. 2006. Evidence for a role of the cellular ND10 protein PML in mediating intrinsic immunity against human cytomegalovirus infections. *J Virol,* 80**,** 8006-18.

TERHUNE, S., TORIGOI, E., MOORMAN, N., SILVA, M., QIAN, Z., SHENK, T. & YU, D. 2007. Human Cytomegalovirus UL38 Protein Blocks Apoptosis. *Journal of Virology,* 81**,** 3109-3123.

TERHUNE, S. S., MOORMAN, N. J., CRISTEA, I. M., SAVARYN, J. P., CUEVAS-BENNETT, C., ROUT, M. P., CHAIT, B. T. & SHENK, T. 2010. Human cytomegalovirus UL29/28 protein interacts with components of the NuRD complex which promote accumulation of immediate-early RNA. *PLoS Pathog,* 6**,** e1000965.

THOMPSON, M. R., KAMINSKI, J. J., KURT-JONES, E. A. & FITZGERALD, K. A. 2011. Pattern Recognition Receptors and the Innate Immune Response to Viral Infection. *Viruses,* 3**,** 920-940.

TIRABASSI, R. S. & PLOEGH, H. L. 2002. The Human Cytomegalovirus US8 Glycoprotein Binds to Major Histocompatibility Complex Class I Products. *Journal of Virology,* 76**,** 6832-6835.

TOMASEC, P., BRAUD, V. M., RICKARDS, C., POWELL, M. B., MCSHARRY, B. P., GADOLA, S., CERUNDOLO, V., BORYSIEWICZ, L. K., MCMICHAEL, A. J. & WILKINSON, G. W. G. 2000. Surface Expression of HLA-E, an Inhibitor of Natural Killer Cells, Enhanced by Human Cytomegalovirus gpUL40. *Science,* 287**,** 1031-1033.

TOMASEC, P., WANG, E. C. Y., DAVISON, A. J., VOJTESEK, B., ARMSTRONG, M., GRIFFIN, C., MCSHARRY, B. P., MORRIS, R. J., LLEWELLYN-LACEY, S., RICKARDS, C., NOMOTO, A., SINZGER, C. & WILKINSON, G. W. G. 2005. Downregulation of natural killer cell–activating ligand CD155 by human cytomegalovirus UL141. *Nature immunology,* 6**,** 181-188.

TOMAZIN, R., BONAME, J., HEGDE, N. R., LEWINSOHN, D. M., ALTSCHULER, Y., JONES, T. R., CRESSWELL, P., NELSON, J. A., RIDDELL, S. R. & JOHNSON, D. C. 1999. Cytomegalovirus US2 destroys two components of the MHC class II pathway, preventing recognition by CD4+ T cells. *Nat Med,* 5**,** 1039-43.

TOOZE, J., HOLLINSHEAD, M., REIS, B., RADSAK, K. & KERN, H. 1993. Progeny vaccinia and human cytomegalovirus particles utilize early endosomal cisternae for their envelopes. *Eur J Cell Biol,* 60**,** 163-78.

TOWLER, J. C. 2007. *Transcriptome activity of human cytomegalovirus (strain Merlin) in fibroblasts, epithelial cells and astrocytes.* PhD thesis Thesis, University of Glasgow.

TOWLER, J. C., EBRAHIMI, B., LANE, B., DAVISON, A. J. & DARGAN, D. J. 2012. Human cytomegalovirus transcriptome activity differs during replication in human fibroblast, epithelial and astrocyte cell lines. *Journal of General Virology,* 93**,** 1046-1058.

TRGOVCICH, J., CEBULLA, C., ZIMMERMAN, P. & SEDMAK, D. D. 2006. Human Cytomegalovirus Protein pp71 Disrupts Major Histocompatibility Complex Class I Cell Surface Expression. *Journal of Virology,* 80**,** 951-963.

TROMBETTA, E. S. & HELENIUS, A. 1998. Lectins as chaperones in glycoprotein folding. *Curr Opin Struct Biol,* 8**,** 587-92.

VALÉS-GÓMEZ, M. & REYBURN, H. T. 2006. Intracellular Trafficking of the HCMV Immunoevasin UL16 Depends on Elements Present in both its Cytoplasmic and Transmembrane Domains. *Journal of Molecular Biology,* 363**,** 908-917.

VANARSDALL, A. L. & JOHNSON, D. C. 2012. Human cytomegalovirus entry into cells. *Current Opinion in Virology,* 2**,** 37-42.

VANARSDALL, A. L., WISNER, T. W., LEI, H., KAZLAUSKAS, A. & JOHNSON, D. C. 2012. PDGF receptor-alpha does not promote HCMV entry into epithelial and endothelial cells but increased quantities stimulate entry by an abnormal pathway. *PLoS Pathog,* 8**,** e1002905.

VARANI, S., CEDERARV, M., FELD, S., TAMMIK, C., FRASCAROLI, G., LANDINI, M. P. & SODERBERG-NAUCLER, C. 2007. Human cytomegalovirus differentially controls B cell and T cell responses through effects on plasmacytoid dendritic cells. *J Immunol,* 179**,** 7767-76.

VARNUM, S. M., STREBLOW, D. N., MONROE, M. E., SMITH, P., AUBERRY, K. J., PAŠA-TOLIĆ, L., WANG, D., CAMP, D. G., RODLAND, K., WILEY, S., BRITT, W., SHENK, T., SMITH, R. D. & NELSON, J. A. 2004. Identification of Proteins in Human Cytomegalovirus (HCMV) Particles: the HCMV Proteome. *Journal of Virology,* 78**,** 10960-10966.

WAGNER, C. S., LJUNGGREN, H.-G. & ACHOUR, A. 2008. Immune Modulation by the Human Cytomegalovirus-Encoded Molecule UL18, a Mystery Yet to Be Solved. *The Journal of Immunology,* 180**,** 19-24.

WALDMAN, W. J., KNIGHT, D. A., HUANG, E. H. & SEDMAK, D. D. 1995. Bidirectional transmission of infectious cytomegalovirus between monocytes and vascular endothelial cells: an in vitro model. *J Infect Dis,* 171**,** 263-72.

WANG, C., ZHANG, X., BIALEK, S. & CANNON, M. J. 2011. Attribution of congenital cytomegalovirus infection to primary versus non-primary maternal infection. *Clin Infect Dis,* 52**,** e11-3.

WANG, D., BRESNAHAN, W. & SHENK, T. 2004. Human cytomegalovirus encodes a highly specific RANTES decoy receptor. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 16642-16647.

WANG, D. & SHENK, T. 2005a. Human cytomegalovirus UL131 open reading frame is required for epithelial cell tropism. *J Virol,* 79**,** 10330-8.

WANG, D. & SHENK, T. 2005b. Human cytomegalovirus virion protein complex required for epithelial and endothelial cell tropism. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 18153-18158.

WANG, E. C., MCSHARRY, B., RETIERE, C., TOMASEC, P., WILLIAMS, S., BORYSIEWICZ, L. K., BRAUD, V. M. & WILKINSON, G. W. 2002. UL40-mediated NK evasion during productive infection with human cytomegalovirus. *Proc Natl Acad Sci U S A,* 99**,** 7570-5.

WANG, E. C. Y., PJECHOVA, M., NIGHTINGALE, K., VLAHAVA, V.-M., PATEL, M., RUCKOVA, E., FORBES, S. K., NOBRE, L., ANTROBUS, R., ROBERTS, D., FIELDING, C. A., SEIRAFIAN, S., DAVIES, J., MURRELL, I., LAU, B., WILKIE, G. S., SUÁREZ, N. M., STANTON, R. J., VOJTESEK, B., DAVISON, A., LEHNER, P. J., WEEKES, M. P., WILKINSON, G. W. G. & TOMASEC, P. 2018. Suppression of costimulation by human cytomegalovirus promotes evasion of cellular immune defenses. *Proceedings of the National Academy of Sciences,* 115**,** 4998-5003.

WANG, J., GUO, Y., WANG, X., ZHAO, R. & WANG, Y. 2017. Modulation of global SUMOylation by Kaposi's sarcoma-associated herpesvirus and its effects on viral gene expression. *J Med Virol,* 89**,** 2011-2019.

WANG, W., YU, P., ZHANG, P., SHI, Y., BU, H. & ZHANG, L. 2008. The infection of human primary cells and cell lines by human cytomegalovirus: New tropism and new reservoirs for HCMV. *Virus Research,* 131**,** 160-169.

WANG, X., HUONG, S. M., CHIU, M. L., RAAB-TRAUB, N. & HUANG, E. S. 2003. Epidermal growth factor receptor is a cellular receptor for human cytomegalovirus. *Nature,* 424**,** 456-61.

WARMING, S., COSTANTINO, N., COURT, D. L., JENKINS, N. A. & COPELAND, N. G. 2005. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Research,* 33**,** e36-e36.

WATFORD, W. T., MORIGUCHI, M., MORINOBU, A. & O'SHEA, J. J. 2003. The biology of IL-12: coordinating innate and adaptive immune responses. *Cytokine Growth Factor Rev,* 14**,** 361-8.

WATSON, R. T. & PESSIN, J. E. 2001. Transmembrane domain length determines intracellular membrane compartment localization of syntaxins 3, 4, and 5. *Am J Physiol Cell Physiol,* 281**,** C215-23.

WEEKES, M. P., TOMASEC, P., HUTTLIN, EDWARD L., FIELDING, CERI A., NUSINOW, D., STANTON, RICHARD J., WANG, EDDIE C., AICHELER, R., MURRELL, I., WILKINSON, GAVIN W., LEHNER, PAUL J. & GYGI, STEVEN P. 2014. Quantitative Temporal Viromics: An Approach to Investigate Host-Pathogen Interaction. *Cell,* 157**,** 1460-1472.

WELCH, A. R., WOODS, A. S., MCNALLY, L. M., COTTER, R. J. & GIBSON, W. 1991. A herpesvirus maturational proteinase, assemblin: identification of its gene, putative active site domain, and cleavage site. *Proceedings of the National Academy of Sciences,* 88**,** 10792-10796.

WELTE, S. A., SINZGER, C., LUTZ, S. Z., SINGH-JASUJA, H., SAMPAIO, K. L., EKNIGK, U., RAMMENSEE, H. G. & STEINLE, A. 2003. Selective intracellular retention of virally induced NKG2D ligands by the human cytomegalovirus UL16 glycoprotein. *Eur J Immunol,* 33**,** 194-203.

WESTSTRATE, M. W., GEELEN, J. L. & VAN DER NOORDAA, J. 1980. Human cytomegalovirus DNA: physical maps for restriction endonucleases BglII, hindIII and XbaI. *J Gen Virol,* 49**,** 1-21.

WHITE, E. A. & SPECTOR, D. H. 2007. Early viral gene expression and function. *In:* ARVIN, A., CAMPADELLI-FIUME, G., MOCARSKI, E., MOORE, P. S., ROIZMAN, B., WHITLEY, R. &

YAMANISHI, K. (eds.) *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.* Cambridge: Cambridge University Press.

WHITLEY, R. J. 1996. Herpesviruses. *In:* BARON, S. (ed.) *Medical Microbiology.* 4th ed. Galveston (TX) University of Texas Medical Branch at Galveston.

WHITMIRE, J. K. 2011. Induction and function of virus-specific CD4+ T cell responses. *Virology,* 411**,** 216-228.

WILKIE, G. S., DAVISON, A. J., KERR, K., STIDWORTHY, M. F., REDROBE, S., STEINBACH, F., DASTJERDI, A. & DENK, D. 2014. First Fatality Associated with Elephant Endotheliotropic Herpesvirus 5 in an Asian Elephant: Pathological Findings and Complete Viral Genome Sequence. 4**,** 6299.

WILKINSON, G. W., KELLY, C., SINCLAIR, J. H. & RICKARDS, C. 1998. Disruption of PML-associated nuclear bodies mediated by the human cytomegalovirus major immediate early gene product. *Journal of General Virology,* 79**,** 1233-1245.

WILKINSON, G. W. G., AKRIGG, A. & GREENAWAY, P. J. 1984. Transcription of the immediate early genes of human cytomegalovirus strain AD169. *Virus Research,* 1**,** 101-116.

WILKINSON, G. W. G., DAVISON, A. J., TOMASEC, P., FIELDING, C. A., AICHELER, R., MURRELL, I., SEIRAFIAN, S., WANG, E. C. Y., WEEKES, M., LEHNER, P. J., WILKIE, G. S. & STANTON, R. J. 2015. Human cytomegalovirus: taking the strain. *Medical Microbiology and Immunology,* 204**,** 273-284.

WILKINSON, G. W. G., TOMASEC, P., STANTON, R. J., ARMSTRONG, M., PROD'HOMME, V., AICHELER, R., MCSHARRY, B. P., RICKARDS, C. R., COCHRANE, D., LLEWELLYN-LACEY, S., WANG, E. C. Y., GRIFFIN, C. A. & DAVISON, A. J. 2008. Modulation of natural killer cells by human cytomegalovirus. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology,* 41**,** 206-212.

WILLE, P. T., KNOCHE, A. J., NELSON, J. A., JARVIS, M. A. & JOHNSON, D. C. 2010. A human cytomegalovirus gO-null mutant fails to incorporate gH/gL into the virion envelope and is unable to enter fibroblasts and epithelial and endothelial cells. *J Virol,* 84**,** 2585-96.

WILLE, P. T., WISNER, T. W., RYCKMAN, B. & JOHNSON, D. C. 2013. Human cytomegalovirus (HCMV) glycoprotein gB promotes virus entry in trans acting as the viral fusion protein rather than as a receptor-binding protein. *MBio,* 4**,** e00332-13.

WILLS, M. R., ASHIRU, O., REEVES, M. B., OKECHA, G., TROWSDALE, J., TOMASEC, P., WILKINSON, G. W., SINCLAIR, J. & SISSONS, J. G. 2005. Human cytomegalovirus encodes an MHC class I-like molecule (UL142) that functions to inhibit NK cell lysis. *J Immunol,* 175**,** 7457-65.

WOOD, L. J., BAXTER, M. K., PLAFKER, S. M. & GIBSON, W. 1997. Human cytomegalovirus capsid assembly protein precursor (pUL80.5) interacts with itself and with the major capsid protein (pUL86) through two different domains. *J Virol,* 71**,** 179-90.

WOODHALL, D. L., GROVES, I. J., REEVES, M. B., WILKINSON, G. & SINCLAIR, J. H. 2006. Human Daxx-mediated repression of human cytomegalovirus gene expression correlates with a repressive chromatin structure around the major immediate early promoter. *J Biol Chem,* 281**,** 37652-60.

WU, J., SONG, Y., BAKKER, A. B., BAUER, S., SPIES, T., LANIER, L. L. & PHILLIPS, J. H. 1999. An activating immunoreceptor complex formed by NKG2D and DAP10. *Science,* 285**,** 730-2.

WUSSOW, F., CHIUPPESI, F., MARTINEZ, J., CAMPO, J., JOHNSON, E., FLECHSIG, C., NEWELL, M., TRAN, E., ORTIZ, J., LA ROSA, C., HERRMANN, A., LONGMATE, J., CHAKRABORTY, R., BARRY, P. A. & DIAMOND, D. J. 2014. Human Cytomegalovirus Vaccine Based on the Envelope gH/gL Pentamer Complex. *PLOS Pathogens,* 10**,** e1004524.

WYATT, J., J., S., R., L. & E., A. 1950. Generalized cytomegalic inclusion disease. *J Pediat***,** 271-294.

YAMAJI, T., NISHIKAWA, K. & HANADA, K. 2010. Transmembrane BAX inhibitor motif containing (TMBIM) family proteins perturbs a trans-Golgi network enzyme, Gb3 synthase, and reduces Gb3 biosynthesis. *J Biol Chem,* 285**,** 35505-18.

YU, D., SILVA, M. C. & SHENK, T. 2003. Functional map of human cytomegalovirus AD169 defined by global mutational analysis. *Proc Natl Acad Sci U S A,* 100**,** 12396-401.

YU, D., SMITH, G. A., ENQUIST, L. W. & SHENK, T. 2002. Construction of a self-excisable bacterial artificial chromosome containing the human cytomegalovirus genome and mutagenesis of the diploid TRL/IRL13 gene. *J Virol,* 76**,** 2316-28.

YU, X., TRANG, P., SHAH, S., ATANASOV, I., KIM, Y.-H., BAI, Y., ZHOU, Z. H. & LIU, F. 2005. Dissecting human cytomegalovirus gene function and capsid maturation by ribozyme targeting and electron cryomicroscopy. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 7103-7108.

YUROCHKO, A. D., HWANG, E. S., RASMUSSEN, L., KEAY, S., PEREIRA, L. & HUANG, E. S. 1997. The human cytomegalovirus UL55 (gB) and UL75 (gH) glycoprotein ligands initiate the rapid activation of Sp1 and NF-kappaB during infection. *J Virol,* 71**,** 5051-9.

ZHANG, G., RAGHAVAN, B., KOTUR, M., CHEATHAM, J., SEDMAK, D., COOK, C., WALDMAN, J. & TRGOVCICH, J. 2007. Antisense transcription in the human cytomegalovirus transcriptome. *J Virol,* 81.

ZHAO, G.-N., ZHANG, P., GONG, J., ZHANG, X.-J., WANG, P.-X., YIN, M., JIANG, Z., SHEN, L.-J., JI, Y.-X., TONG, J., WANG, Y., WEI, Q.-F., WANG, Y., ZHU, X.-Y., ZHANG, X., FANG, J., XIE, Q., SHE, Z.-G., WANG, Z., HUANG, Z. & LI, H. 2017. Tmbim1 is a multivesicular body regulator that protects against non-alcoholic fatty liver disease in mice and monkeys by targeting the lysosomal degradation of Tlr4. *Nature Medicine,* 23**,** 742.

ZHOU, M., YU, Q., WECHSLER, A. & RYCKMAN, B. J. 2013. Comparative Analysis of gO Isoforms Reveals that Strains of Human Cytomegalovirus Differ in the Ratio of gH/gL/gO and gH/gL/UL128-131 in the Virion Envelope. *Journal of Virology,* 87**,** 9680-9690.

ZHU, H., SHEN, Y. & SHENK, T. 1995. Human cytomegalovirus IE1 and IE2 proteins block apoptosis. *Journal of Virology,* 69**,** 7960-7970.

ZIEHR, B., VINCENT, H. A. & MOORMAN, N. J. 2016. Human Cytomegalovirus pTRS1 and pIRS1 Antagonize Protein Kinase R To Facilitate Virus Replication. *J Virol,* 90**,** 3839-48.

ZISCHKE, J., MAMARELI, P., POKOYSKI, C., GABAEV, I., BUYNY, S., JACOBS, R., FALK, C. S., LOCHNER, M., SPARWASSER, T., SCHULZ, T. F. & KAY-FEDOROV, P. C. 2017. The human cytomegalovirus glycoprotein pUL11 acts via CD45 to induce T cell IL-10 secretion. *PLOS Pathogens,* 13**,** e1006454.