An Analysis of Human Cytomegalovirus Gene Usage

A thesis submitted in candidature for the degree of
DOCTOR OF PHILOSOPHY

by

Sepehr Seirafian

December 2012

Institute of Infection and Immunity

Cardiff University

Declaration

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed …… (candidate) Date……

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of ……(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed …… (candidate) Date ……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references.

Signed …… (candidate) Date ……

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying 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 for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Graduate Development Committee.**

Signed …… (candidate) Date ……

Acknowledgements

I would like to express my sincerest gratitude to my supervisors Prof. Gavin Wilkinson and Dr Peter Tomasec for the opportunity to study a PhD, and for their constant encouragement, patience and endeavours.

I wholeheartedly thank my colleagues at the Institute of Infection and Immunity for making this PhD a pleasurable experience. Special thanks to Dr James Davies and Dr Richard Stanton for their ever-present support in the laboratory, Dr Rebecca Aicheler for the NK cell inhibition assays, Dr Simone Cuff for the generation of polyclonal sera, and Dr Eddie Wang, Dr Ceri Fielding, Dr Jason Twohig and Dr Stephen Clark for sharing valuable insights. Thanks also to Daniel Sugrue for his contribution to immunofluorescence imaging, William Perks for assistance in data processing, Dr Carole Rickards for help until her retirement, and Mrs Dawn Roberts for assistance since her appointment.

I wish to thank my parents for their relentless support throughout my studies at University. Special thanks must also go to my grandfather and A.Wesolowska for showing keen interest and kind concern for my progress in writing this thesis.

Summary

HCMV encodes a plethora of immune-modulating functions, many of which have yet to be assigned to specific genes. In prospect of performing high-throughput screens to identify and characterise such functions, a library of recombinant adenoviruses (RAds) each encoding a V5 epitope-tagged HCMV protein was generated. Protein expression was validated and characterised for the vast majority of RAds by western blot and immunofluorescence.

HCMV has been reported to both upregulate cell surface expression of Fas, and render cells resistant to Fas-mediated killing. This thesis demonstrated that Fas levels are markedly reduced at the surface of HCMV-infected cells as an early function that persists through the late phase. Screening a panel of HCMV deletion mutants eliminated 83 genes as not required for Fas downregulation, while screening the RAd library did not identify any single HCMV gene as being sufficient for this function.

Deep sequencing of the HCMV transcriptome recently led to the identification of UL150A as a novel protein-coding gene. To test this prediction, UL150A was tagged within the strain Merlin genome. UL150A was shown to encode multiple protein products, and be expressed with early and late kinetics.

In a screen of the RAd library, gpUL4 was observed to be secreted from cells. To investigate this function in the context of HCMV infection, an epitope-tag was inserted at the 3'-end of the UL4 gene in the strain Merlin genome. Tagged gpUL4 was secreted from cells infected with strain Merlin. Secreted gpUL4 was more heavily glycosylated, and produced in greater abundance than its intracellular counterpart late in infection. Active secretion would be consistent with gpUL4 acting as a virokine, cytokine or cytokine/chemokine-binding protein. gpUL4 purified from supernatants of Merlin- or RAd-UL4-infected cells inhibited NK cell degranulation. Furthermore, gpUL4 did not co-purify with virus particles, indicating it is unlikely to be a virion component.

Contents

List of tables

List of figures

Abbreviations

°C	Degrees Centigrade
µg	Microgram
µl	Microlitre
µM	Micromolar
µm	Micrometer
AC	Assembly compartment
Ad	Adenovirus
ADPRT	Adenosine diphosphate ribosyl transferase
AIDS	Acquired immunodeficiency syndrome
ANT	Adenine nucleotide translocator
ARP	Actin-related protein
AST	Antisense transcripts
ATP	Adenosine triphosphate
BAC	Bacterial artificial chromosome
BAL	Bronchoalveolar lavage
BLAST	Basic local alignment search tool
BP	Base pair
BPP	Basic phosphoprotein
BSA	Bovine serum albumin
cAMP	Cyclic Adenosine monophosphate
cDNA	Complementary DNA
CBP	CREB-binding protein
CCMV	Chimpanzee cytomegalovirus
CD	Cluster of differentiation

CF	Complement fixation
CIDV	Cytomegalic inclusion disease virus
CKBP	Chemokine-binding protein
CMV	Cytomegalovirus
CNS	Central nervous system
CPE	Cytopathic effect
CR	Coding region
CREB	cAMP response element-binding protein
CRS	*cis* repression sequence
CTL	Cytotoxic T lymphocyte
DAB	3, 3’ diaminobenzidine
DAPI	4, 6’ diamino-2-phenylindole
DAXX	Death-domain associated protein
DB	Dense body
DMEM	Dulbecco’s modified Eagle’s medium
DMSO	Dymethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide triphosphate
DOG	2-deoxygalactose
DTT	Dithiothreitol
DURP	dUTPase-related protein
EDTA	Ethylenediaminetetraacetic acid
E.Coli	Escherichia coli
EGFP	Enhanced green fluorescent protein
EGFR	Epidermal growth factor receptor
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
ER	Endoplasmic reticulum

ERF	Ets-2 repressor factor
EV	Ebola virus
FCS	Foetal calf serum
FSC	Forward scatter
*gal*K	Galaktokinase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
gB	Glycoprotein B
gCI	Glycoprotein complex I
gCII	Glycoprotein complex II
gCIII	Glycoprotein complex III
GFP	Green fluorescent protein
gH	Glycoprotein H
gL	Glycoprotein L
gM	Glycoprotein M
gN	Glycoprotein N
gO	Glycoprotein O
gp	Glycoprotein
GPCR	G protein-coupled receptor
GRIM-19	Gene associated with retinoid-IFN-induced mortality 19
HBV	Hepatitis B virus
HCMV	Human cytomegalovirus
hCAR	Human Coxsackie adenovirus receptor
HCM	Hydrochloric acid
HDAC	Histone deacetylase
hDAXX	Human death-domain associated protein
HHV	Human herpesvirus
HLA	Human leukocyte antigen
HMVEC	Human microvascular endothelial cell
HIV	Human immunodeficiency virus

HSV	Herpes simplex virus
HF	Human fibroblasts
HFFF	Human foetal foreskin fibroblast
HMWP	High molecular weight protein
HRP	Horseradish peroxidise
HSPG	Heparin sulphate proteoglycans
hTERT	Human telomerase reverse transcriptases
IE	Immediate early
Ig	Immunoglobulin
IL	Interleukin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRL	Inverted repeat long
IRS	Inverted repeat short
kDa	Kilodalton
LA	Latex agglutination
LB	Lysogeny broth
LIR-1	Leukocyte inhibitory receptor-1
LMP	Lower matrix protein
LPS	Lipopolysaccharide
LUNA	Latent undefined nuclear antigen
mAb	Monoclonal antibody
MCP	Major capsid protein
mCP	Minor capsid protein
mC-BP	Minor capsid protein-binding protein
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
MOI	Multiplicity of infection

mRNA	Messenger RNA
miRNA	MicroRNA
MV-LAP	Myxoma virus leukaemia –associated protein
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NF-κB	Nuclear factor kappaB
NIEP	Noninfectious enveloped particles
NK cell	Natural killer cell
NKG2D	Natural killer group 2, member D
NLS	Nuclear localisation signal
NNT	noncoding nonoverlapping transcripts
NP40	Nonyl phenoxypolyethoxylethanol-40
NT2D1	NTera-S clone D1 (teratocarcinoma cell line)
OAS	Oligoadenylate synthetase
OD	Optical density
ORF	Open reading frame
oriLyt	Lytic origin of DNA replication
p	Phosphate
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDGFR	Platelet-derived growth factor receptor
pfu	Plaque-forming unit
p.i.	Post infection
PKR	Protein kinase R
PML	Promyelocytic leukemia
PolyA	Polyadenylation signal
pp	Phosphoprotein
Q-PCR	Polymerase chain reaction

RACE	Rapid amplification of cDNA ends
RAd	Recombinant adenovirus
RANTES	Regulated and normal T cell expressed and secreted
RB	Retinoblastoma
RFP	Red fluorescent protein
RID	Receptor internalisation and degradation
RL	Repeat long
RNA	Ribonucleic acid
RNase	Ribonuclease
RPE	Retinal pigment epithelium
PRV	pseudorabies virus
rRNA	Ribosomal RNA
RSV	Respiratory syncytial virus
RT-PCR	Reverse transcription PCR
SAS	Superacceptor site
SCP	Small capsid protein
Scrapin	Surface cellular receptor abductor protein
SILAC	Stable isotope labelling by amino acids in cell culture
SLAM	Signalling lymphocyte-activation molecule
SSC	Side scatter
TAE	Tris-acetate-EDTA
TAP	Transporter associated with antigen processing
TBP	TATA-binding protein
TBST	Tris-buffered saline Tween-20
tet	Tetracycline
TFIIB	Transcription factor IIB
THP-1	Human monocytic leukemia cell line
TLR	Toll-like receptor
TNF	Tumour necrosis family

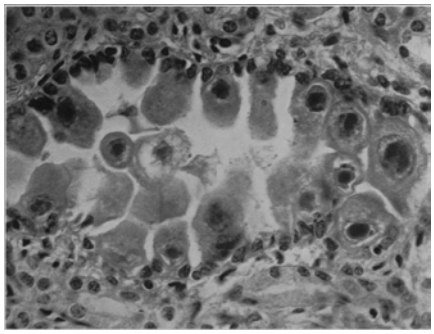
TNFR	Tumour necrosis family receptor
TRAIL	TNF-related apoptosis-inducing ligand
TREx	Tetracycline repressor protein
TRL	Terminal repeat long
tRNA	Transfer RNA
TRS	Terminal repeat short
UL	Unique long
ULBP	UL16-binding protein
UMP	Upper matrix protein
uORF	Upstream open reading frame
US	Unique short
UV	Ultraviolet
vICA	Viral inhibitor of caspase-8 activation
vMIA	Viral mitochondrion-localised inhibitor of apoptosis
VSV	Vesicular stomatitis
X-gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside

1-Introduction

1.1. History of isolation of HCMV

In 1881, Ribbert first noted the presence of 'protozoan-like' structures in the kidney of a still born infant. These findings were not published until 1904 (Ribbert, 1904), whilst earlier that year, Jesionek and Kiolemenoglou had reported on the observation of these unusual large cells, in the kidneys of an 8 month old fetus suffering from hereditary syphilis (Jesionek and Kiolemenoglou, 1904). These cells were later concluded to be protozoa, and were named *Endameba mortinatalium* (Smith and Weidman, 1910, Smith and Weidman, 1914). Goodpasture in 1921 provided evidence that such structures were not protozoa, but rather abnormal cytomorphosed cells, to which he gave the name cytomegalia (large cell) (Goodpasture and Talbot, 1921). Example images of these cytomegalic cells are shown in Figure 1.1. By 1950, conclusive evidence had been obtained that the aforementioned cellular morphology was caused by viral infection (Wyatt et al., 1950). This included identification of similar cytomegalic cells containing nuclear and cytoplasmic inclusions, in the cells of other animals such as guinea pigs, hamsters, mice, and monkeys, and their constant association with a serially transmissible infectious agent able to cause lethal disease (Wyatt et al., 1950). The term cytomegalic inclusion disease virus (CIDV) was proposed for the causative agent of diseases known as 'inclusion disease', and 'salivary gland disease' (Wyatt et al., 1950). This virus was first isolated from the adenoids of children undergoing tonsillectomy-adenoidectomy surgery (Rowe et al., 1953) whilst its serial propagation following inoculation of fibroblast cultures was for the first time reported by Smith (Smith, 1956). Soon after, 2 other cases of successful recovery of the virus were reported (Rowe et al., 1956, Weller et al., 1957). As it was recognised that the laboratories had isolated different strains of the same virus (e.g. Davies and AD169), and to simplify the nomenclature, the name 'cytomegalovirus' was proposed and adopted (Weller et al., 1960).

a) **adapted from Farber and Wolbach 1932**



b) **adapted from Wyatt et al. 1950**

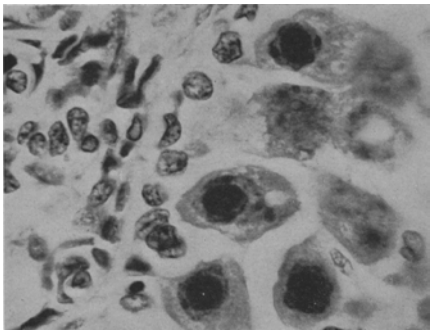


Figure 1.1: **'Protozoan-like' cytomegalic cells**
'protozoan-like' structures were actually HCMV infected cells: large inclusion-laden cells containing swollen nuclei lining the kidney tubule.

Human Cytomegalovirus (HCMV) infection *in vivo* gives rise to large, conspicuous oval-shaped cells containing characteristic intranuclear and cytoplasmic granules or 'inclusion bodies', thought to be respectively areas of virus DNA replication and assembly (Iwasaki et al., 1973).

In 1961, HCMV was classified as a herpes virus, based on evidence that it was a DNA virus, and that its size closely resembled that of previously identified herpes viruses (Andrewes et al., 1961). Morphological findings achieved by electron microscopy supported the above proposal, and HCMV has now been established as the prototype β-herpesvirus (Smith and Rasmussen, 1963, Spaete et al., 1994).

1.2. HCMV Genome

1.2.1. Genome organization

HCMV has the largest genome (235 kbp) among all human viruses discovered to date (Dolan et al., 2004). The linear double-stranded HCMV genome is of the E class of herpesvirus genomes, and is formed by the covalent bonding of two sequences of varying length, known as unique long (UL) and unique short (US) (Roizman, 1980). As in other herpesviruses of the E class genome, the HCMV UL and US segments are flanked by the inverted repeats TRL-IRL, and IRS-TRS respectively (Fleckenstein et al., 1982). The TRL-IRL, and IRS-TRS inverted repeats are also referred to as *b* and *c* respectively. At the termini of the genome, there are direct repeats of a short sequence, termed the *a* sequence, which is also present in an inverted state at the UL-US junction (Spaete and Mocarski, 1985). The overall genome configuration therefore is *a*-TRL-UL-IRL-*a*'-IRS-US-TRS-*a*, or *ab*-UL-*b'a'c'*-US-*ca*, where the prime designation signifies sequence inversion (Spaete and Mocarski, 1985). A diagrammatic representation of the genome is given in Figure 1.2. The *a* sequence can vary in copy number depending on the strain and passage of the virus (Zaia et al., 1990), and contains two highly conserved cis-acting elements called *pac*1 and *pac*2, adjacent to which the genome is cleaved prior to packaging (Kemble and Mocarski, 1989). The 'a' sequence has also been demonstrated to be indispensable for inversion, amplification, and circularization of the genome (Spaete and Mocarski, 1985). Following cleavage, genome circularisation

occurs through ligation of single unpaired bases present at each 3' terminus of the HCMV DNA (Tamashiro and Spector, 1986). Recombination between the terminal, and internal repeat sequences gives rise to inversion of the two unique segments relative to each other, resulting in four equimolar isomers of the HCMV genome (Weststrate et al., 1980). It is unknown if isomerisation of the HCMV genome brings about functional consequences, although isomer specific long-range splicing may occur (Gatherer et al., 2011).

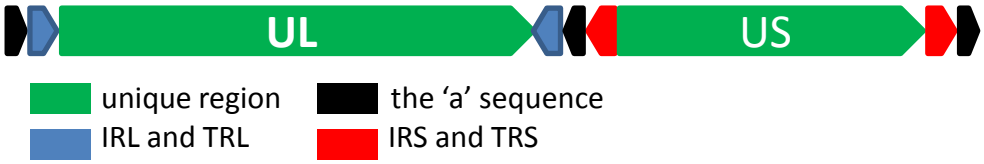


Figure 1.2: **Organization of the HCMV genome**
Unique long (UL) and unique short (US) regions are flanked by inverted repeats named TRL and IRL (also named b and b’), and IRS and TRS (also named c and c’) respectively. The ‘a’ sequence is present at the termini of the genome, and also at the U_L-U_S junction in an inverted state. The direction of the arrow indicates the direction of the sequence. Recombination between a terminal ‘a’ sequence and the inverted ‘a’ sequence at the UL-US junction gives rise to four equimolar genomic isomers differing in the orientation of the unique sequences. The above isomer is conventionally chosen to depict the HCMV genome

1.2.2. HCMV Strains

Several different HCMV strains have been studied in detail. The general genome organization is remarkably similar between viruses isolated from different patients, tissues, and geographical locations. HCMV strains are often referred to as high or low passage based on the extent of culture following isolation from clinical samples.

High passage strains have been extensively passaged in fibroblasts following isolation from clinical material, often with the purpose of attenuating the virus for vaccine development. Adaptation to continuous *in vitro* passaging of HCMV isolates resulted in significant genetic change, including loss and duplication of large segments of the genome, and the accumulation of several frame-shift mutations (Dolan et al., 2004). These genetic alterations manifested themselves as changes in the biological characteristics of the virus, such as accelerated growth kinetics, and a narrowing of the range of cell types the virus is able to replicate in (Cha et al., 1996, Sinzger et al., 2000). Indeed an exceptional property of HCMV is its capacity to infect a wide range of cell types *in vivo*. The loss of endotheliotropism in passaged HCMV strains has been linked with the UL128 locus genes (UL128, UL130, and UL131A), at least one of which is invariably mutated in these strains (Dolan et al., 2004, Akter et al., 2003). All three of the above genes are known to be essential for HCMV growth in endothelial cells (Hahn et al., 2004), and their function may be disadvantageous to HCMV growth in fibroblasts (Adler et al., 2006).

High passage strains, and in particular AD169 and Towne were the subjects of the majority of HCMV research before the discovery of their major genetic deletions (Cha et al., 1996). In AD169, 20 ULb' genes (UL133-UL150 inclusive) have been replaced by an inverted duplication of sequences in the RL region, and genes RL5A, RL13 and UL131A are frame-shifted (Dolan et al., 2004). In addition the UL36 gene is inactivated by a single amino acid substitution (Skaletskaya et al., 2001). Towne lacks ULb' genes UL147-UL150, and is mutated in genes RL13 and UL130 (Murphy et al., 2003). Regrettably, these defective strains have continued to be widely used as research materials, despite the availability of more complete strains. This is partly due to their efficient replication in fibroblasts, yielding higher titers of virus, and thereby facilitating experimentations.

Low passage strains contain more complete versions of the HCMV genome. Six such strains of HCMV have been cloned in Bacterial Artificial Chromosomes (BACs) and sequenced: FIX, TR, PH, Toledo (Murphy et al., 2003), TB40/E (Sinzger et al., 2008) and Merlin (Stanton et al., 2010). The first four of the above underwent an undefined number of passages before being cloned into BACs. TB40/E was passaged 5 times in fibroblasts and 22 times in endothelial cells (ECs). Merlin was sequenced and BAC-cloned after 3 and 5 passages respectively (Dolan et al., 2004, Stanton et al., 2010). It is noteworthy that Toledo has undergone an inversion encompassing 14.4 kb of the UL/*b*' region, resulting in loss of function of the UL128 gene (Cha et al., 1996, Akter et al., 2003). The FIX BAC originated from virus taken from a pregnant mother with primary HCMV infection (Hahn et al., 2002, Grazia Revello et al., 2001). TR was isolated from the eye of an AIDS patient (Smith et al., 1998), whilst PH and TB40/E were removed from transplant patients with HCMV disease (Rice et al., 1984, Fish et al., 1995, Sinzger et al., 1999). Toledo (Quinnan Jr et al., 1984), and later Merlin (Quinnan Jr et al., 1984, Tomasec et al., 2000) were isolated from the urine of two congenitally infected children.

Importantly, during the process of BAC cloning, alterations were made to the genomes of most of the viruses in order to accommodate essential BAC vector sequences. In Towne and Toledo the BAC sequences replaced US1-US11 and US2-US11 respectively. PH and FIX both lack the IRS1-US6 region, and TR is missing US2-US5 (Murphy et al., 2003).

Merlin is the only clinical strain BAC-cloned in its entirety (Stanton et al., 2010). Unlike other HCMV BAC clones, the original clinical material from which Merlin was isolated is available and can be accessed when required (Dolan et al., 2004). With the exception of single point mutations in the RL13 and UL128 genes which have since been repaired, strain Merlin contains the full complement of HCMV genes. In addition Merlin is designated the prototype HCMV genome sequence by the National Centre for Biotechnology Information (GenBank AY446894; RefSeq NC_006273).

1.2.3. Number of genes

1.2.3.1. HCMV gene usage

Strain AD169 was the first HCMV genome to be completely sequenced, and the original annotation identified and designated 208 ORFs (Chee et al., 1990). In this annotation, an ORF was 'designated' if it encoded at least 100 amino acids, and did not overlap a longer ORF by more than 60% of its length. These criteria were based on the observation that in previously studied herpesvirus genomes few overlapping ORFs had been noted, and that the longest overlapping ORF was normally the coding ORF (Chee et al., 1990). It is noteworthy that 14 of the 208 strain AD169 ORFs were present in a large repeated sequence in the left hand end of the UL region and therefore diploid. Later, 19 additional ORFs were discovered in Toledo, through sequence analysis of the UL/*b*' region (Cha et al., 1996). The duplication in the UL region is thought to have arisen as compensation for the UL/*b*' deletion, since encapsidation requires a near full length genome (Prichard et al., 2001). When Murphy et al. reported the full sequencing of BAC cloned strains Toledo, FIX, PH, TR , Towne and AD169 (Murphy et al., 2003), as many as 252 ORFs were designated in low passage strains. To reassess the HCMV gene layout, a comparison of the AD169 and Toledo genomes was carried out against the closest known relative of HCMV, chimpanzee cytomegalovirus (CCMV) (Davison et al., 2003b). Results from other functional and bioinformatic studies, and evidence for presence of HCMV-specific genes were also considered. Through this analysis, 51 ORFs previously annotated by Chee et al (Chee et al., 1990) were eliminated due to their lack of CCMV orthologues, whilst 10 new ORFs were delineated. With the exception of HCMV-specific ORFs, this re-annotation predicted that wildtype HCMV contains 164-167 protein-coding genes (Davison et al., 2003b). Deep sequencing of the HCMV transcriptome during productive lytic infection, combined with extensive analysis of HCMV genomic sequences has resulted in some amendment to HCMV gene usage, with the most recent update predicting 170 protein-coding genes (Gatherer et al., 2011). Figure 1.3 displays the current annotation of genes for strain Merlin. The possibility that the number of translated genes may vary depending on the cell type investigated, between lytic and latent infection, or indeed different stages of lytic infection cannot be neglected. Therefore failure to detect certain transcripts in deep sequencing of the transcriptome of cells undergoing lytic

infection cannot be regarded as absolute evidence that their respective ORFs are non-coding. For instance, in monocytes which are regarded as a site of HCMV latency, an antisense transcript to the UL81 and UL82 genes, named UL81-UL82ast, and its resulting protein product, latent undefined nuclear antigen LUNA have been observed (Bego et al., 2005). Later, LUNA was shown to be expressed *in vivo*, and antibodies to the LUNA protein were detected in the blood of seropositve, but not seronegative individuals (Bego et al., 2011).

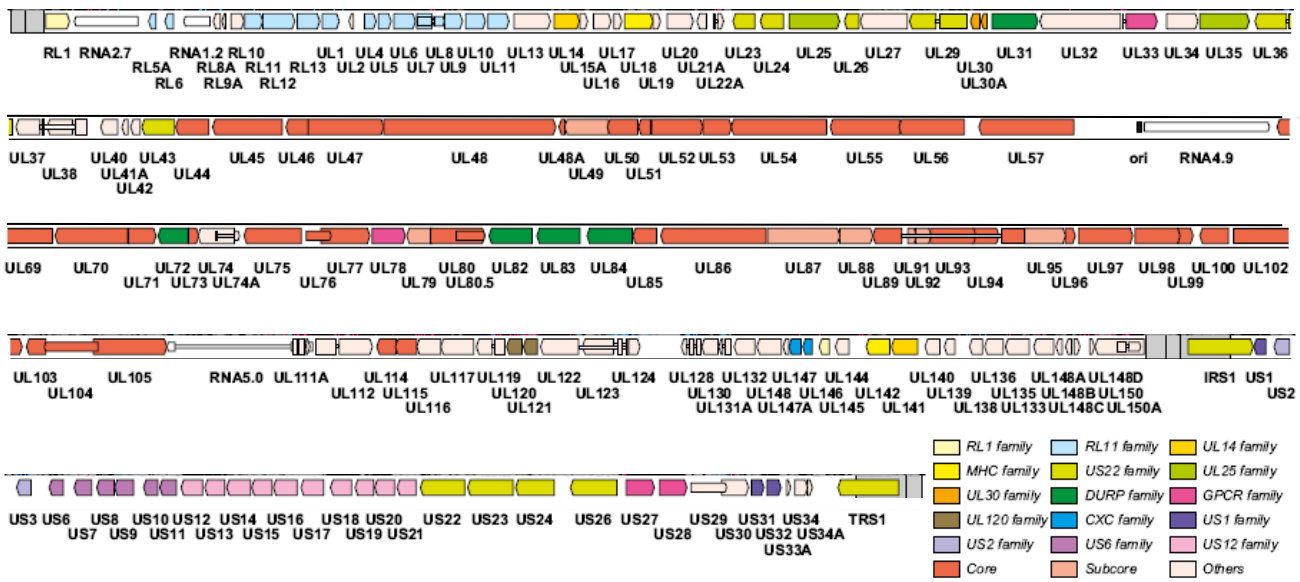


Figure 1.3 (adapted from Gatherer et al. 2011):

Diagram of the wildtype HCMV genome (Merlin strain). The linear double-stranded HCMV genome is presented in five segments. Genes are shown as coloured arrows, with the direction indicating sense strand. Gene members of the same family are matched in colour (see key), and core genes are indicated in brown. The repeat sequences at the termini of the genome and at the UL-US junction are colured grey. Non-coding non-overlapping transcripts (NNTs) are shown by white arrows, and introns by narrow bars within genes; RNA5.0 is both an intron and an NNT. Origin of replication (ori) is presented as a black bar.

1.2.3.2. Nomenclature of HCMV genes and proteins

In the seminal report of the AD169 sequence, Chee et al. proposed an alphanumeric nomenclature of HCMV ORFs, based on their position in the UL, US or repeat regions that provides the basis for the formal nomenclature used today (Chee et al., 1990, Spaete et al., 1994). However, use of various systems in naming HCMV proteins has been somewhat confusing. Multiple nomenclatures have been adopted for HCMV proteins on the basis of their physiochemical properties, predicted position in the virion in case of structural proteins, their temporal expression, function, and homology with other herpes virus proteins, in particular HSV (Spaete et al., 1994). In two separate workshops in 1993, an attempt was made to introduce a consensus nomenclature, and it was agreed that the alphanumeric system described by Chee et al would be extended for HCMV protein products in the following way: the prefix 'p' would be utilised to denote proteins not known to be phosphorylated or glycosylated, whilst 'pp' and 'gp' would precede the ORF name for phosphorylated, and glycosylated proteins respectively (Spaete et al., 1994).

1.2.3.3. Non-coding RNA molecules

Noncoding RNAs account for the majority of transcription from the HCMV genome, and have been divided into 3 main classes (Gatherer et al., 2011, Zhang et al., 2007). The most abundant of these are the noncoding nonoverlapping transcripts (NNT) which originate from regions of the genome largely void of protein coding ORFs, and carry 3' polyA signals (Gatherer et al., 2011). The expression of 4 NNTs has so far been described: RNA2.7, RNA1.2, RNA5.0, and RNA 4.9, where numbers indicate size of the transcript in kb (see Figure 1.3). RNA2.7 (β2.7) was identified as the most abundantly transcribed gene in early stage HCMV infection (Spector, 1996, Greenaway and Wilkinson, 1987), and is the best studied of all HCMV NNTs. It is present in a single copy in low passage isolates, but the laboratory strains AD169 and Towne became diploid for the β2.7 gene following expansion of their TRL regions (described above). The longest ORF in strain AD169 β2.7 is 170 amino acid (aa) long (Greenaway, Wilkinson 1987). This ORF can be translated in a cell free assay (Bergamini et al., 1998), but McSharry and co-workers showed it was not conserved across all strains, and that β2.7 was dispensable for replication of strain AD169 *in vitro* (McSharry et al., 2003). Although reporter genes are extremely efficiently expressed when inserted in place of,

(or within) β2.7 (Spaete and Mocarski, 1987), translation of any natural short putative β2.7 ORFs within the context of HCMV infection has yet to be described, leading to current predictions it is indeed an NNT. More recent findings assigning a function to RNA2.7 are described later (section 1.7.1). RNA1.2, also encoded in the TRL region, was found to be a highly expressed transcript, but similarly to RNA2.7, its translation into protein has only been reported *in vitro* (Hutchinson and Tocci, 1986). RNA5.0 is transcribed with immediate early kinetics (Plachter et al., 1988), and is the intronic sequence delimited by the two exons that give rise to the spliced transcript RNA1.1 in late stage infection (Kulesza and Shenk, 2004). RNA1.1 is not classed as an NNT as it is formed by the splicing together of 2 exons positioned either side of the RNA 5.0 intron, yet peculiarly, the 2 molecules have never been detected in a single primary transcript (Kulesza and Shenk, 2004). RNA4.9 was recently identified, and like RNA5.0, is of unknown function (Gatherer et al., 2011). The above NNTs represent two thirds of the total transcribed RNA at 72 hours p.i., with RNA2.7 accounting for over three quarters of that amount (Gatherer et al., 2011).

The second most abundant class of noncoding RNA transcripts are the antisense transcripts (AST), which result from transcription of the DNA strand antiparallel to a coding region (CR) (Gatherer et al., 2011). These were first revealed through sequence analysis of cDNA clones of the HCMV transcriptome (Zhang et al., 2007). ASTs complementary to most CRs have been detected, and are generally present in lower amounts than their sense strand counterparts (Gatherer et al., 2011). Like NNTs, these are polyadenylated, and their potential to encode protein cannot be ruled out. Sequence conservation and complex splicing observed in some ASTs suggests they may serve functions, with post-transcriptional regulation of the sense transcript being the most obvious candidate.

The third major group of HCMV noncoding RNA transcripts are microRNAs (miRNAs). Potentially, miRNAs provide the virus with a powerful way of exercising temporal and spatial gene regulation. Obvious benefits of such regulation include their low biological production cost, and the avoidance of immunogenic detection (Dölken et al., 2009). The vast majority of virally encoded miRNAs have been discovered in herpesviruses (Grey et al., 2007), 14 of which have been characterised in HCMV. These are named

according to the name of the gene they oppose, or in the case of those originating from noncoding regions, the nearest upstream gene on the same strand.

The first 9 HCMV miRNAs were identified by a computer program that searched for potential miRNA precursors based on sequence and predicted secondary structure. Subsequently, sequencing of cloned small RNA molecules from lytically infected primary human fibroblasts (HF) confirmed these predictions (Pfeffer et al., 2005). Additionally, cloning and sequencing of 18-24 nucleotide RNA molecules from HCMV infected HF cells identified 3 potential miRNAs (Dunn et al., 2005). *In silico* analysis of hypothetical pre-miRNA transcripts giving rise to these molecules supported these findings, and their expression by HCMV was confirmed by northern blot. Grey et al. used a bioinformatics approach to predict stem-loop structures conserved between HCMV and CCMV (Grey et al., 2005). In this study, expression of the predicted miRNAs during HCMV infection was also assessed by northern blot analysis, and 2 novel HCMV miRNAs were discovered.

HCMV miRNAs are expressed throughout the genome from antisense, and noncoding regions, and some are known to target multiple viral, or both viral and cellular mRNAs (Stern-Ginossar et al., 2007b, Grey et al., 2007). Nevertheless, the targets and functions of most HCMV miRNAs remain unknown. The best understood HCMV miRNAs is miR-UL112, which targets multiple viral and cellular transcripts. Grey et al. used a bioinformatics program to search for potential miR-UL112 binding sites within 3' untranslated regions (UTRs) conserved between HCMV and CCMV (Grey et al., 2007). 14 candidates were identified, all of which were cloned into plasmids downstream of a luciferase reporter construct. These were then co-transfected into 293 cells, together with plasmid-based clones of the predicted coding region for miR-UL112. IE72, UL120/121, and UL112/113 were validated as miR-UL112 targets, as luciferase expression was specifically inhibited in the presence of their 3' UTRs. The functional significance of IE72 inhibition by miR-UL112 is still unclear, but the authors proposed that since IE72 expression promotes acute viral replication, its negative regulation by miR-UL112 may contribute to the maintenance of latent or persistent infection. In another study, algorithmic prediction of miRNA target sites identified the major histocompatibility complex class I-related chain B (MICB) as a potential target for miR-UL112 (Stern-Ginossar et al., 2007b). MICB is a stress-induced ligand of the NK

activatory receptor NKG2D, and is important in the NK response against virus-infected cells. MICB expression was downregulated by miR-UL112 in the context of viral infection, leading to resistance against NK-killing. This represented a novel mechanism of immune-evasion by HCMV. Later, it was demonstrated that the UL114 gene too is targeted by miR-UL112, reducing its activity as uracil DNA glycosylase, but with minimal effect on virus growth (Stern-Ginossar et al., 2009).

It is difficult to conceive that the majority of noncoding RNAs are not functionally important. We have only relatively recently become aware of their existence and their host and viral RNA target sequences can be difficult to identify.

1.2.4. Essential and non-essential genes

Biochemical, genetic, and bioinformatic studies have identified the function of many HCMV genes. Despite this, several HCMV genes have yet to be assigned specific functions. In a systematic analysis of the genome, Dunn et al. produced deletion mutants of the 162 genes that the Towne strain was thought to encode based on the annotation of the strain AD169 genome by Chee et al.1990 (Dunn et al., 2003b). 10 non-essential genes were replaced in the process of BAC cloning the virus (US1-US12), whilst the remaining 152 deletions were made in separate viruses. The ability of the mutant viruses to replicate was then tested in each of the resulting mutants. 45 ORFs were found to be indispensible for virus replication in fibroblasts, the majority of which are conserved at least amongst β-herpesviruses. This is in contrast to the non-essential ORFs, which were HCMV-specific in about 70% of cases. Figure 1.4 shows the results of this study in diagrammatic form, where genes have been subdivided into groups based on the growth properties of mutant viruses. HCMV genes that are nonessential for virus replication *in vitro* are nevertheless expected to play important roles *in vivo*. A group have been found to be important in determining cell tropism, evidenced by differential infection and growth in endothelial, epithelial and fibroblastic cells. A substantial number of genes are known to direct immunomodulation in the host. Intriguingly, the study by Dunn identified 4 mutants that enhanced growth, suggesting they act as growth suppressors (Dunn et al., 2003b). Moreover DNA microarray analysis of the HCMV transcriptome in different cell types has revealed significant cell type dependent differences in gene expression (Towler et al., 2012).

HCMV genes essential for virus replication in fibroblast cultures include some major structural proteins, surface glycoproteins required for entry, and those acting in DNA replication. The majority of essential HCMV genes are clustered together in the central region of the genome, whilst the majority of non-essential genes tend to occupy peripheral regions (Dunn et al., 2003b).

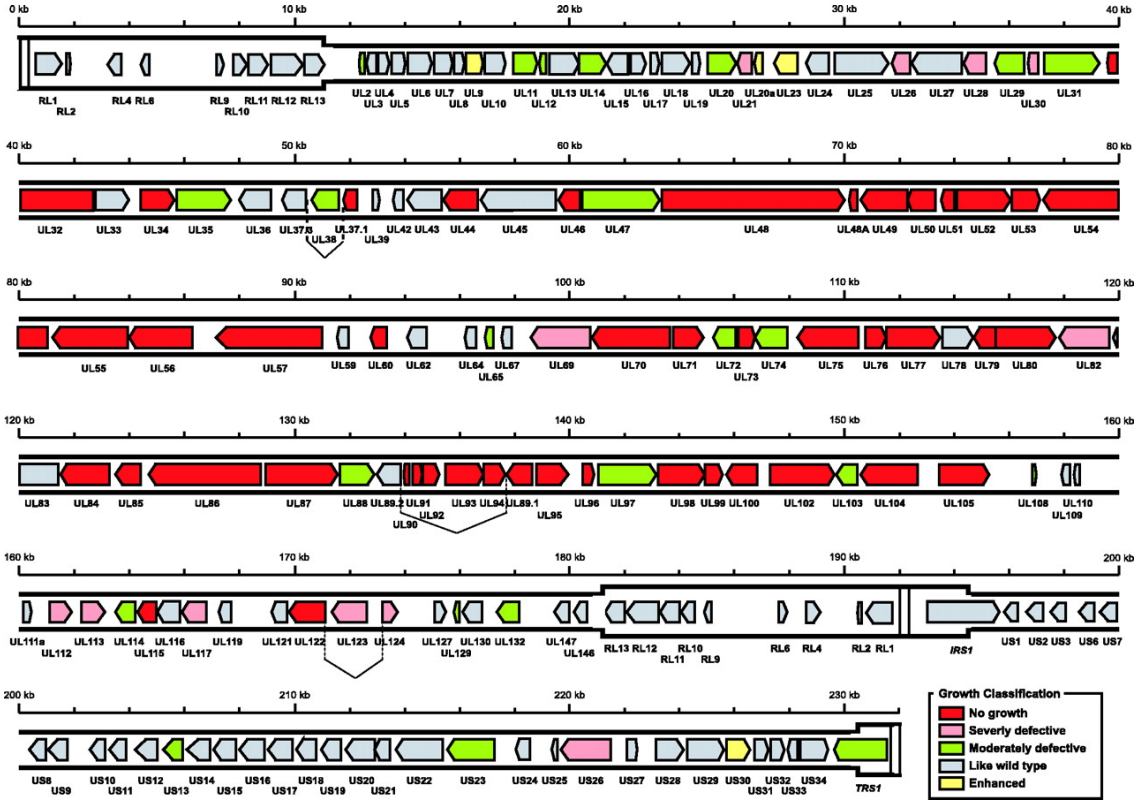


Figure 1.4 **(adapted from Dunn et al. 2003. Copyright 2003, The National Academy of Sciences): Effect of deleting individual HCMV genes on virus growth**

HCMV strain Towne genes colour coded on the basis of the growth properties displayed by their respective deletion mutants in fibroblasts.

1.2.5. Gene families

Based on sequence homology, non-essential HCMV genes have been categorised into 13 families, with the possible exception of UL84 which may be an essential gene (Dunn et al., 2003b, Gatherer et al., 2011, Sarisky and Hayward, 1996). Table 1.1 lists the different gene families and their members, with a brief description of their functions. Members of a family often lie close together in the genome, and can exhibit related functions (Chee et al., 1990, Davison and Bhella, 2007, Davison et al., 2003a, Dolan et al., 2004). HCMV gene families are thought to have arisen through gene duplication with divergence over time. Sequence analysis can be used to estimate a timescale of these duplications. The largest and best conserved HCMV gene family is the US22 family, encoding 13 members which display roles in tropism, pathogenesis, and temperance of virus growth (Dunn et al., 2003b).

Not all non-essential HCMV genes are part of gene families however. Certain genes appear to have homologous host counterparts, and are thought to have been incorporated through gene capture (Davison and Bhella, 2007). In addition, HCMV may have gained or donated genes from/to other viruses; e.g. the RL11 family of glycoproteins share an immunoglobulin domain like motif present in some adenoviruses (Davison et al., 2003a).

1.3. Virion structure

1.3.1. Overview of structure

The complete HCMV virion measures about 210-230 nm in diameter, and is comprised of a nucleocapsid, contained within a partially ordered proteinaceous tegument, which in turn is surrounded by an envelope (Gibson, 1981). The first electron microscopy (EM) images of the HCMV virion revealed round, near spherical particles containing a dense centre, surrounded by what was described as a pale halo (Smith, 1959, Stern and Friedmann, 1960). Improved images of virion particles and additional information on the fine structure of the virion were later obtained by negative staining electron microscopy (Smith and Rasmussen, 1963, Wright Jr et al., 1964). Figure 1.5 shows an electron cryomicrograph of 2 adjacent HCMV particles taken at x30,000 magnification.

The HCMV DNA molecule has an Mr of $1.5x10^8$ (Kilpatrick et al., 1976), more than 60 times greater than the total Mr of all virion proteins (Sarov and Abady, 1975). The first estimates of the number of virion proteins were between 20 and 35, (Sarov and Abady, 1975, Kim et al., 1976, Stinski, 1977, Edwards, 1976) ; ranging in size between 22 and 230 kDa (Fiala et al., 1976). Gibson in 1983 produced evidence of the topological position of some of the virion proteins, and recognised the existence of A, B and C capsids described below (Gibson, 1981). He also identified a set of proteins, including pp65, that were localised between the capsid and the envelope, the fraction now known as the tegument (Gibson, 1981). Improved purification techniques, and the availability of the AD169 sequence lead to the recognition of 6 new structural proteins; and their corresponding genes (Baldick and Shenk, 1996). Later, a mass spectrometry approach discovered 24 novel structural proteins, increasing the number of virion proteins to 59 (Varnum et al., 2004). More recently, RL13 and gpUL1 have been identified as envelope glycoproteins (Shikhagaie et al., 2012a, Stanton et al., 2010).

Table 1.1 **(adapted from Davison and Bhella 2007):**

HCMV gene families, their members, and their encoded functions

Gene family	Family members	Functions
RL11	RL5A, RL6, RL11, RL12, RL13, UL1, UL4, UL5, UL6, UL7, UL8, UL9, UL10, UL11	Mostly membrane glycoproteins
UL14	UL14, UL141	Membrane glycoproteins containing an immunoglobulin domain
UL18	UL18, UL142	MHC-I-related membrane glycoproteins
UL25	UL25, UL35	Tegument proteins
GPCR	UL33, UL78, US27, US28	Chemokine receptors
DURP	UL31, UL72, UL82, UL83, UL84	Some are tegument proteins
UL120	UL120, UL121, possibly UL119	Membrane glycoproteins
UL146	UL146, UL147	CXC chemokines
US1	US1, US31, US32	Related to TT virus ORF2
US2	US2, US3	Membrane glycoproteins
US6	US6, US7, US8, US9, US10, US11	Membrane glycoproteins
US12	US12, US13, US14, US15, US16, US17, US18, US19, US20, US21	Multiple transmembrane proteins
US22	UL23, UL24, UL26, UL28, UL29, UL36, UL43, US22, US23, US24, US26, IRS1, TRS1	Tegument proteins

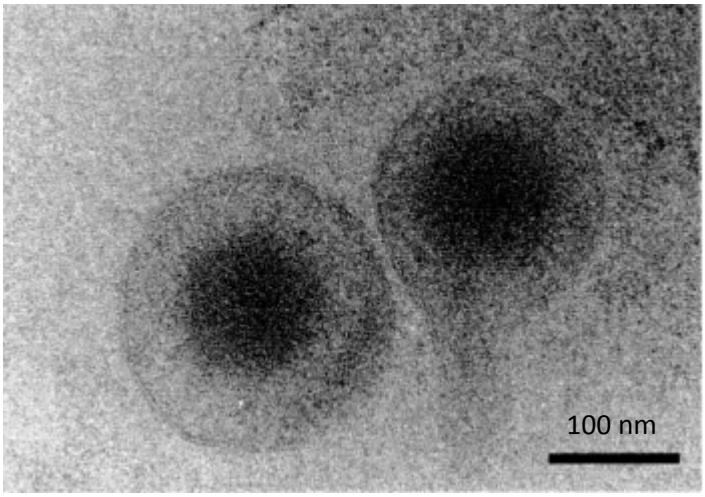


Figure 1.5 (adapted from Bhella et al. 2000):
Electron cryomicrograph of 2 adjacent HCMV particles
The darkly stained central region represents the nucleocapsid. The tegument is formed around the nucleocapsid and is in turn surrounded by the more sharply stained envelope

1.3.2. Nucleocapsid

The nucleocapsid is icosahedral (T=16) with a diameter of 130 nm, and consists of 150 hexameric, and 12 pentameric capsomeres, forming the faces and edges, and the vertices respectively (Chen et al., 1999). Each hexon is made of 6 copies of the major capsid protein (MCP; pUL86) and 6 copies of the small capsid protein (SCP; pUL48A) at the tips, whilst MCP is the sole element in pentomeric capsomeres (Chen et al., 1999). Capsomeres are connected together by a total of 320 triplexes, which are heteromers consisting of 2 copies of the minor capsid protein (mCP; pUL85), and 1 copy of the mCP-binding protein (mC-BP; pUL46) (Chen et al., 1999). Three different species of capsids have been isolated from HCMV infected cells all containing an identical arrangement of the 4 integral capsid proteins (Gibson, 1981). 'A' capsids lack viral DNA and are essentially empty proteinaceous shells (Chen et al., 1999). B capsids also lack DNA, but in addition to the capsid icosahedron they also contain copies of the capsid scaffold proteins pUL80 and pUL80.5 (Chen et al., 1999). It is thought that B capsids undergo DNA encapsidation and lose the internal scaffold proteins to give rise to mature C capsids. C capsids act as precursors of the infectious virion, and are associated with certain tegument proteins. The viral DNA enters the nucleocapsid through a portal complex adjacent to one of the pentomers (Dittmer and Bogner, 2005) and is tightly packaged into regularly arranged duplexes separated only by 23 A^{o} (Zhou et al., 1999, Bhella et al., 2000). Following encapsidation 90% of the internal volume of the capsid is predicted to be occupied by the viral DNA (Yu et al., 2003). Furthermore although all nucleocapsids appear uniform in size, their relative location inside the virion has been shown to vary (Chen et al., 1999).

1.3.3. Tegument

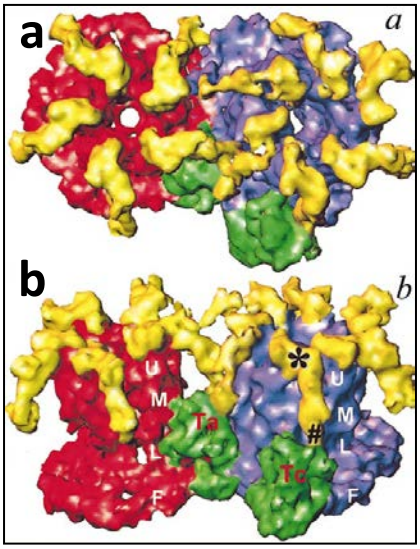
The tegument accounts for approximately 100 nm of the diameter, and 40% of the total protein mass of the virion (Gibson, 1996). Thus far more than 35 HCMV tegument proteins have been identified, several of which are phosphorylated and highly immunogenic (Kalejta, 2008). No common sequence has been identified between tegument proteins that would guide virion constituents into the tegument. Many HCMV tegument proteins are encoded by a cluster of HCMV genes, and this feature has been used to predict additional members of the tegument entity (Mocarski, 1993).

The variation in the relative position of nucleocapsids in the virions, and the flexibility in the size of the tegument suggests most tegument proteins lack icosahedral symmetry and are not in rigid contact with the nucleocapsid (Spaete et al., 1994, Chen et al., 1999). 3D density mapping of the HCMV virion has revealed disconnected densities within the envelope and the majority of the tegument, supporting this notion (Chen et al., 1999). The only icosahedrally structured portion of the tegument appears to be the capsid-proximal region, forming a thin net-like layer of loosely connected densities, which is absent from the B capsids but present in C capsids and the intact virion as evidenced by difference mapping studies. These data suggest the capsid acts as a scaffolding base around which the tegument is sequentially built, explaining why the ordered tegument layer has never been observed independently of the capsid (Chen et al., 1999). It has been proposed that these densities are formed by the basic phosphoprotein (BPP; pp150, UL32), the upper matrix protein (UM; pp71, UL82), and pp28 (UL99). These 3 proteins are present in large quantities in the tegument, and furthermore they have been shown to associate most stringently with the nucleocapsid after detergent treatment of the virions (Chen et al., 1999, Gibson, 1996). Also, immunoelectron microscopy of purified virions suggested a capsid proximal location for pp28 (Landini et al., 1987), whilst pp150 and pp71 have both been shown to be present in approximately equimolar quantities to the MCP (Chen et al., 1999). The ordered tegument consists of 960 identical copies of a cluster formed between these proteins, each one making contact with a triplex, and a penton or hexon. A prediction of the interaction between the nucleocapsid and the ordered tegument is given in Figure 1.6. The other 3 major tegument proteins are the high molecular weight protein (HMWP; pUL48), its binding protein (HMWP-binding protein; pUL47), and the lower matrix protein (LMP; ppUL83; pp65). Despite being the most abundant protein species in the tegument, pp65 is non-essential for virus replication and the production of infectious virus progeny (Schmolke et al., 1995b). This suggests pp65 serves to fill the tegument volume thereby assisting with the assembly of virus particles (Liu and Zhou, 2007). The immune-modulating function assigned to pp65 is discussed in (Section 1.7.3.1.1). pp71, the upper matrix protein (UMP) encoded by UL82 relieves transcriptional repression of the immediate early (IE) genes by degrading DAXX (Bresnahan and Shenk, 2000b). Other tegument proteins include pUL35 and pUL69 (Winkler et al., 1994, Winkler and Stamminger, 1996) which are activators of viral gene

expression, the anti-apoptotic proteins pUL36 and pUL38, and pUL44 and pUL54, the processivity and catalytic subunits of the HCMV DNA polymerase (Kalejta, 2008). pIRS1 and pTRS1 have been demonstrated to be associated with virions and dense bodies, which lack capsids. Treatment of intact HCMV particles with trypsin did not degrade virion pIRS1 and pTRS1, suggesting they are located within, and not outside the envelope, and that they may be tegument proteins. In the presence of UL69, these 2 proteins can promote the expression of reporter genes containing IE gene promoters, and have therefore been implicated in early HCMV gene expression (Romanowski et al., 1997).

Figure 1.6 (adapted from Chen et al. 1999): Capsid proteins interact with the ordered tegument densities

Interactions between capsid and tegument proteins Viewed from top (a) or side (b): Region of capsid shown contains one penton (red), one hexon (blue), and 2 adjacent triplexes labelled Ta and Tc (green). Penton and hexon subunits contain upper (U), middle (M), lower (L), and floor (F) domains. A cluster of tegument densities is shown in yellow, whilst elevated and lower ends of one of the tegument densities are labelled * and # respectively. Tegument densities are bridged over the intercapsomeric space, and are attached by their elevated end to a penton or hexon, and by their lower ends to a triplex which is used as a pier. No direct contact is observed between tegument densities of a single cluster.



1.3.4. Envelope

The HCMV virion envelope is a lipid bilayer derived from the host intracellular membranes, containing at least 25 HCMV glycoproteins (Vanarsdall and Johnson, 2012), and numerous host proteins such as CD55, CD59, and annexin II (Grundy et al., 1987, Wright et al., 1995). It is thought the host proteins in the viral envelope help to induce host cellular responses in combination with some viral encoded G-protein coupled receptors (Compton et al., 2003, Zhu et al., 1998).

Envelope glycoproteins are important targets of the host neutralising antibody response and their interaction with the host cell is indispensable for virus entry (Mocarski, 2001). Many HCMV virion glycoproteins have been shown to form disulfide-linked complexes. The best characterised of these include glycoprotein complex I (gCI), gCII, and gCIII (Britt and Boppana, 2004). gCI consists of 55 kDa and 116 kDa gB isoforms, gM-gN heterodimers give rise to gCII, and gCIII is formed by interaction between gH, gL and gO. Of the aforementioned glycoproteins, gB, gM, gN, gH and gL are essential for virus replication (Hobom et al., 2000). In wildtype virus, rather than complexing with gO, gH and gL bind the UL128 locus (128L) proteins, UL128, UL130 and UL131A. This pentomer is essential for entry into epithelial, endothelial and certain myeloid cells, but prevents virus growth in fibroblasts (Stanton et al., 2010). Upon passaging of HCMV in fibroblast cultures, UL128L genes rapidly mutate, leading to loss of the above pentamer. Viruses with the mutated UL128L can then infect fibroblasts through the formation of gCII. This represents an example of tropism-specific complex formation at the HCMV envelope.

gCII (gM/gN) is the most abundant glycoprotein complex at the virus envelope. gM is predicted to have 7 transmembrane domains, and undergo a single N-linked glycosylation (Britt and Boppana, 2004), whilst gN is thought to be a type I glycoprotein containing O-linked modification. gM and gN form a complex in the ER, and are both essential for virus replication. Between different HCMV strains, very little sequence variability has been observed in the case of gM, in contrast with gN, which is highly variable in its sequence (Britt and Boppana, 2004, Dal Monte et al., 2001). This variability suggests gN is under selective pressure from the immune system, and it has

been proposed that glycosylation protects gN from antibody binding (Wei et al., 2003, Britt and Boppana, 2004).

gB encoded by the UL55 gene is the best conserved, and one of the most immunogenic HCMV glycoproteins (Cranage et al., 1986). For these reasons gB has been the prime subunit vaccine candidate (Spaete et al., 1994). Cleavage of the 150 kDa type I gB precursor by the Golgi localised protease furin, gives rise to two cleavage products measuring 93 and 55 kDa. These cleavage products are linked by disulfide bonding to form a heterodimer at the envelope and undergo phosphorylation, and N- and O – linked glycosylation (Britt and Boppana, 2004).

1.3.5. Other virion components

Evidence has been obtained of the selective incorporation of cellular, and virus-encoded RNAs in the HCMV virion (Greijer et al., 2000). The presence of HCMV-encoded RNAs was first noted in virion purification experiments, and was evidenced by gene array experiments (Bresnahan and Shenk, 2000a). The abundance of these transcripts in the virion has been shown to correlate with their intracellular concentration, and their packaging into the virion likely involves non-specific interaction with structural proteins (Terhune et al., 2004). Virus-encoded RNAs in the virion include the transcript encoding the UL22A secreted protein (R27080), and the 1.2 kb, 2.7 kb and 5kb NNTs (Bresnahan and Shenk, 2000a). GAPDH, B-actin and cyclin G1 are amongst the cellular RNA transcripts detected in the virion (Terhune et al., 2004). Although it has been demonstrated that HCMV consistently incorporates the above transcripts into the virion, it is likely that at least some of these RNAs serve no particular structural function, and that their packaging is nonspecific, resulting either from accidental affinity to certain virion proteins, or their high concentration in the cell (Greijer et al., 2000). One obvious advantage of carrying polyA containing RNAs in the infectious particle would be their immediate expression in the infected cell prior to the initiation of transcription from the viral genome.

Over 70 cellular proteins have been reported in HCMV virions, including structural proteins, enzymes and chaperon proteins, most of which are likely contaminants (Varnum et al., 2004). Some cellular proteins found in the virion, including actin-related protein (ARP) and polyamines may act to stabilise the virion. ARP is thought to

reside in the capsid-distal portion of the tegument, whilst positively charged polyamines may facilitate packaging and carriage of viral DNA by neutralising its highly negative charge (Gibson and Roizman, 1971).

1.3.6. Dense bodies and NIEPs

In addition to complete virions, 2 other types of enveloped particles are recovered from the culture of HCMV-infected cells. Dense bodies (DBs) are spherical particles comprised largely of the pp65 tegument protein encased within the envelope, but lacking the nucleocapsid. The diameter of DBs is highly heterogeneous and is typically much larger than that of virions (250-600 nm) (Sarov and Abady, 1975). Noninfectious enveloped particles (NIEPs) also lack genetic material, but in contrast to DBs, these are almost identical in protein architecture to the infectious virion, except they contain B-capsids (Irmiere and Gibson, 1983). Virions, DBs and NIEPs are distinguishable under the electron microscope, and can be separated by density gradient centrifugation (Irmiere and Gibson, 1983, Talbot and Almeida, 1977). Figure 1.7 compares electromicrographs of virions and DBs, and a Coomassie stained gel of proteins in virions, DBs and NIEPs is shown in Figure 1.8. The production of DBs and NIEPs by infected cells can exceed that of infectious virions by up to 20 times; yet It is unknown if they represent products of aberrant virus assembly, or serve specific functions (Gibson, 1996, Steven and Spear, 1997). It has been suggested that DBs and NIEPs may serve as decoy targets of the immune system (Gibson, 1996, Steven and Spear, 1997).

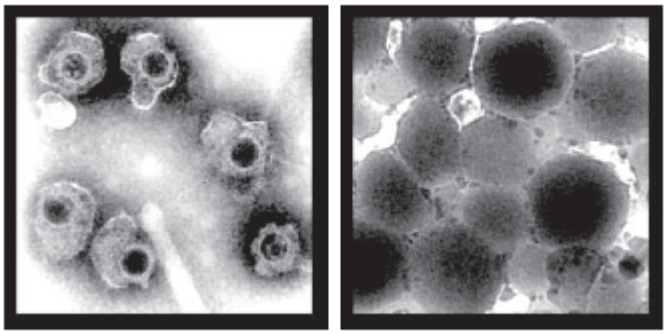
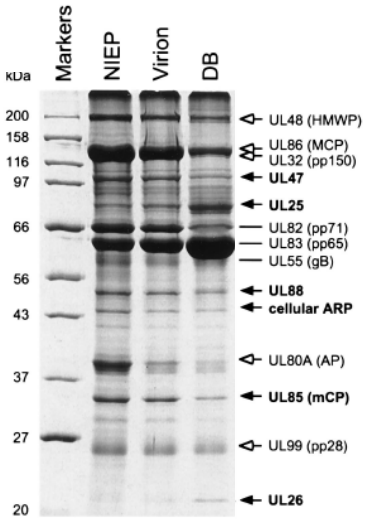


Figure 1.7 (adapted from Varnum et al. 2004): Electron micrographs of HCMV virions and dense body (DB) particles

Magnification, x8,400. The heterogenous and larger size of DBs relative to virions are noticeable.

Figure 1.8 (adapted from Baldick Jr and Shenk 1996):
Visualization of proteins in HCMV particles by Coomassie staining
Coomassie-stained gel of virions, NIEPs and DBs purified by gradient centrifugation. Near-identical protein profiles are noticeable between NIEPs and virions, with only notable difference being the concentration of the capsid assembly protein (AP) which is present in the B-type capsids of NIEPs and not in the C-type capsids of mature virions. Also of note is the scarcity of capsid proteins, and the high concentration of tegument proteins (particularly pp65) in DBs.



1.4. Life cycle and gene expression

1.4.1. Tropism

Efficient HCMV replication is restricted to human and chimpanzee cells *in vitro*, but investigations of *in vivo* infections have detected viral replication in an extraordinarily wide range of human tissues and cell types. From a purely biological perspective, cell tropism is defined by the ability of the virus to infect a particular cell type.

In vivo, epithelial, endothelial, fibroblast, and smooth muscle cells are thought to be some of the most efficient sites of HCMV replication (Sinzger et al., 1995). Infection of hematopoietic and endothelial cells facilitates the spread of virus within the host, and mucosal epithelial cell infection is important in inter-host transmission. Cell types thought not to support HCMV infection include lymphocytes and polymorphonuclear leukocytes. Early reports of immediate early (IE) gene expression in lymphocytes (Rice et al., 1984) have not been validated (Sinzger et al., 1995). In polymorphonuclear leukocytes, despite virus uptake, replication is soon aborted (Grefte et al., 1994, Sinzger et al., 1996). Clinical manifestations of HCMV infection appear most frequently in the liver, gastrointestinal tract, lung, retina, and the brain (Grefte et al., 1994, Sinzger et al., 1996). Within these organs, hepatocytes, alveolar epithelial cells, and neuronal cells are often most heavily infected.

The majority of cell culture research on HCMV is carried out on skin and lung fibroblasts. HCMV infection in these cells is by far the most productive, and is the standard method of isolation and propagation of the virus (Sinzger et al., 1995). As discussed in (Section 1.2.2) extensive passaging of HCMV in fibroblasts causes the production of increasingly non-endothelial and non-epithelial tropic virus, whereas propagation in endothelial and epithelial cells maintains the broad cell specificity of the virus. It is plausible to consider that similar cell type dependent adaptations may also occur *in vivo* (Waldman et al., 1991).

Apart from the UL128L genes discussed in (Section 1.2.2), other genes have also been discovered to play important roles in the determination of HCMV tropism. Dunn et al identified 68 ORFs in the Towne strain, the deletion of each of which had no significant effect (<10 fold) on the growth rate of the virus. 15 of these mutants were also tested

in cell cultures of clinically relevant retinal pigment epithelial (RPE), and human microvascular endothelial cells (HMVEC). Two mutants were identified, one which conferred a severe growth defect in HMVECs (UL24), and another which resulted in a growth defect RPE cells (UL64) (Dunn et al., 2003b). Conversely, viruses carrying UL10 and US16 mutations grew to 100 and 500-fold higher titers in RPE and HMVEC cells respectively, compared to HFFs (Dunn et al., 2003b). These were the first findings of tropism determining genes outside of the UL128L. More recently, RL13 was reported to severely impair virus replication in cultures of fibroblasts and epithelial cells (Stanton et al., 2010). Growth suppression by specific viral genes and in specific cell types may be important *in vivo*, for example in moderating viral load in order to optimise transmission.

1.4.2. Entry

Entry of the HCMV particle into the cell is a complex, multistep process that initiates with the sequential binding of viral envelope glycoproteins to cell surface receptors. Direct fusion between the virus envelope and the plasma membrane at neutral pH promotes entry into the cell (Compton et al., 1992). Membrane fusion in RPE and endothelial cells differs from this model however, as it requires acidification (Ryckman et al., 2006).

A comprehensive list of HCMV genes essential for entry is not available, but gB, gH and gL are certainly required (Isaacson and Compton, 2009, Bowman et al., 2011, Vanarsdall and Johnson, 2012). The UL128L products are essential for entry into epithelial and endothelial cells (Akter et al., 2003, Dolan et al., 2004, Hahn et al., 2004, Patrone et al., 2005), and gO-mutants give rise to undersized plaques in fibroblasts (Wille et al., 2010). gB is involved in initial attachment and subsequent fusion, whilst gM/gN also act in attachment. gH is a fusion receptor, which together with its chaperone protein gL, either binds gO to give rise to gCIII, or with the UL128L products to form a pentamer conferring tropism in endothelial and epithelial cells (Ryckman et al., 2006).

At least *in vitro*, cell surface heparin sulphate proteoglycans (HSPGs) are essential for HCMV entry into the cell (Compton et al., 1993). The association between HSPGs and envelope glycoproteins gB and gM is thought to allow later events by stabilizing the

virion at the cell surface (Compton et al., 1993). The quest for other HCMV receptors though has yielded little success, and HSPGs remain the only molecules known to be indispensible for virus entry. Three integrin heterodimers (α2β1, α6β1, αVβ3) have been demonstrated to be HCMV receptors, as virion entry levels are reduced in their absence, and restored by their subsequent addition (Feire et al., 2004). Numerous other candidates have been proposed and later discounted upon further investigation, including β_2-microglobulin, CD13, and MHC Class I molecules (E. Murphy 2008). Annexin II, and an uncharacterised 92.5 kDa cell surface protein may enhance binding and entry (Adlish et al., 1990, Keay and Baldwin, 1995), whilst promising initial reports on the role of epidermal growth factor receptor (EGFR) (Wang et al., 2003) have since been contradicted (Isaacson et al., 2007). Furthermore EGFRs cannot be considered universal HCMV receptors as they are not expressed by all HCMV permissive cells (Real et al., 1986). Soroceanu et al. reported that platelet-derived growth factor-α receptor (PDGFR-α) was phosphorylated by HCMV in a wide range of cell types, and that its genetic deletion or functional blockage prevented HCMV infection in fibroblasts (Soroceanu et al., 2008). Furthermore addition of PDGFR-α to knockout cells restored infectivity to these cells, and gB was demonstrated to interact directly with PDGFR-α. Strong evidence therefore exists for the important role this receptor plays in HCMV entry, although myeloid cells which are important sites of latency and reactivation (Section 1.4.4) do not express PDGFR-α (Chan et al., 2009, Inaba et al., 1993).

1.4.3. Productive infection/ lytic cycle

1.4.3.1. Overview

The lytic cycle starts from the time of virus entry into the cell, and lasts until exit of virus progeny. The HCMV lytic cycle is relatively slow, and new virus progeny are not released until 48-72 hours p.i. (Rice et al., 1984). In HFs, peak levels of virus release are reached at about 72-96 hours p.i., and productive infection can continue for several days, until cell death (Stinski, 1983).

After virus entry, the HCMV genome must be transported into the nucleus in order for viral transcription to occur. Gene expression then occurs in a tightly regulated cascade, which with the aid of metabolic inhibitors has operationally been divided into immediate early (IE), early and late stages (Wathen and Stinski, 1982). IE viral

transcription occurs from restricted regions of the genome, independently of *de novo* viral protein synthesis. Early genes expression requires de novo viral protein synthesis, occurs before viral DNA replication, and is dependent on the presence of IE proteins. Early gene products are in turn required for late RNAs synthesis, which begins after the onset of viral DNA replication (Wathen and Stinski, 1982, Demarchi et al., 1980).

1.4.3.2. Nuclear transport of viral DNA

Once inside the cell, the majority of tegument proteins are released into the cytoplasm (Ogawa-Goto et al., 2003). The nucleocapsid together with some tightly associated tegument proteins, then deliver the genome to the nucleus using the intracellular transport machinery, and through the nuclear pore complex (Greber and Way, 2006). A functioning microtubule network is essential for the nuclear transport of viral DNA, as its chemical disruption inhibits HCMV gene expression (Ogawa-Goto et al., 2003). The tegument proteins that accompany the nucleocapsid on its journey to the nucleus include pp150, pUL47, pUL48, pp71, and pp65. The nonessential pp65 protein contains a nuclear localisation signal (NLS) (Schmolke et al., 1995a), and accumulates in the nucleus within minutes of virus entry into the cell, where it suppresses the induction of many interferon-responsive, and proinflammatory chemokine transcripts (Browne and Shenk, 2003). pUL47 and pUL48 form a complex known to be important in virus entry, and injecting viral DNA into the nucleus (Bechtel and Shenk, 2002), while an interaction between pp150 and the microtubule network promotes intracellular transport of capsid particles (AuCoin et al., 2006). In addition pp71 brings about IE genes expression by degrading the transcriptional repressor DAXX (Bresnahan and Shenk, 2000b, Woodhall et al., 2006).

1.4.3.3. Immediate early (IE or α) genes

At least four distinct regions in the HCMV genome are expressed with IE kinetics; these are schematically represented in Figure 1.9. IE gene products promote efficient virus replication and counter host intrinsic, innate, and adaptive immune responses (Dimitropoulou et al., 2010).

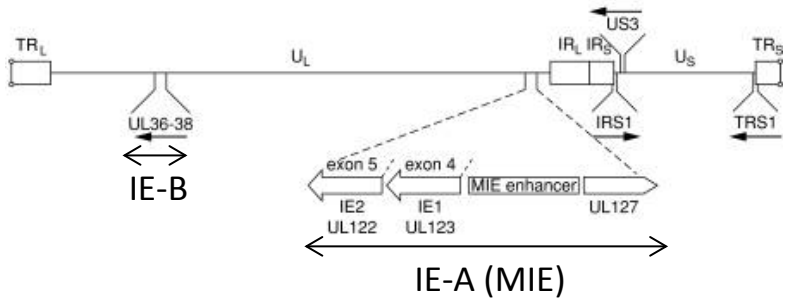


Figure 1.9 (adapted from Stinski and Meier 2007):
Schematic representation of IE regions of the HCMV genome:
Immediate early (IE) regions are indicated with arrows.

Expression from the major immediate early (MIE) gene is under the control of a complex promoter region containing a TATA box, an imperfect dyad symmetry element, and a strong enhancer comprising 17, 18, 19 and 21 bp repeat motifs (Boshart et al., 1985, Thomsen et al., 1984). The MIE region contains MIE genes IE1(UL123) and IE2 (UL122), which are alternatively spliced from a single transcript, and give rise to the well characterised IE72 and IE86 proteins respectively, although other smaller IE2 gene products are also expressed (Akrigg et al., 1985, Wilkinson et al., 1984). IE72 and IE86 are localised in the nucleus, and play key roles in regulating viral and cellular gene expression. IE72 is not essential for *in vitro* virus replication, but its absence results in severe growth defect at low multiplicity of infection (Greaves and Mocarski, 1998, Mocarski et al., 1996). Numerous functions have been described for IE72. on its own, IE72 is a strong activator of the MIE promoter (Cherrington and Mocarski, 1989) and a weak transactivator of several cellular and viral promoters (Wilkinson et al., 1998). IE72 does not directly bind DNA, but physically associates with several cellular transcription factors and nuclear proteins (Nevels et al., 2004). Direct binding to the CCAAT-box binding protein (CTF1) and E2F-1 activates the human DNA polymerase α and dihydrofolate reductase genes, both essential for DNA replication (Hayhurst et al., 1995, Margolis et al., 1995). IE72 targetting of p107 results in the de-repression of E2F-responsive promoters, and the subsequent induction of numerous genes associated with cell cycle progression (Poma et al., 1996). Strong evidence therefore exists for the contribution of IE72 to cell cycle progression, and viral DNA replication.

IE72 co-localises with condensed chromatin during mitosis, and disrupts ND10 (PML-bodies), resulting in the uniform dispersal of PML throughout the nucleus (Kelly et al., 1995, Wilkinson et al., 1998). ND10 are normally seen as spherical intranuclear structures found in association with PML and other cellular proteins including hDAXX, sp100, p53, STAT1 and STAT2, and have been implicated in transcription, chromatin structure, DNA repair and apoptosis (Dimitropoulou et al., 2010). During viral infection, PML acts as a mediator of the interferon-induced antiviral response, and its RNAi knockdown, and that of hDAXX have been shown to enhance HCMV replication (Wang et al., 1998, Tavalai et al., 2006). ND10 are therefore recognised as an intrinsic defence mechanism against viral infection (Dimitropoulou et al., 2010). Moreover, live cell

imaging of an IE1-EGFP fusion protein during infection demonstrated that IE72-mediated disruption of ND10 is not limited to the displacement of PML. STAT2 and sp100 also associated with IE72 on metaphase chromatin although the latter later underwent progressive degradation until almost complete elimination from the nuclei of infected cells. hDAXX was released into the nucleoplasm, this is significant because hDAXX also contributes to the intrinsic antiviral defence described above. Furthermore the presence of IE72 enhances transactivation by IE86 (Malone et al., 1990).

IE86 is a prolific transactivator of viral and cellular genes, and also negatively autoregulates the MIE promoter (Hermiston et al., 1990, Pizzorno et al., 1988, Stenberg et al., 1990). IE86 is essential for HCMV replication, and regulates gene expression in multiple ways (White et al., 2004). Repression of the MIE promoter occurs through direct binding to the *cis* repression sequence (CRS), which is located between the promoter TATA box and the transcription start site (Arlt et al., 1994). This autorepression is concomitant with the recruitment of chromatin-remodelling enzymes by IE86, to establish a repressive chromatin state at the MIE promoter (Reeves et al., 2006). IE86-mediated transactivation of cyclin E is also through direct DNA attachment (Arlt et al., 1994), whilst protein-protein interactions have been reported between IE86 and several transcription factors including Dr1, the TATA-binding protein (TBP), TFIIB, the retinoblastoma protein (RB), CREB, C-Jun, and JunB (Caswell et al., 1996, Caswell et al., 1993, Furnari et al., 1993, Hagemeier et al., 1994, Jupp et al., 1993, Lang et al., 1995, Scully et al., 1995, Sommer et al., 1994). In contrast to its role in promoting a repressed chromatin state at the MIE promoter, IE86 is able to physically interact with the histone acetyltransferase P/CAF, leading to the formation of active chromatin at its responsive promoters (Bryant et al., 2000). This explains at least partly, the capacity of IE86 to act as a promiscuous gene transactivator, and highlights its pleiotropic functionality in gene regulation.

The IE-B region of the HCMV genome gives rise to three known protein products, from genes UL36, UL37 and UL38. Although transactivating roles have been described for these molecules, it is likely their main functions are anti-apoptotic (see Section 1.7.1).

TRS1 and IRS1 proteins are incorporated into the virion tegument, and viruses harbouring mutations in both of these genes show defective growth (Blankenship and

Shenk, 2002). Furthermore at least one of these two genes is required for orilyt-dependent DNA replication (Pari et al., 1993). Together with the MIE gene products, and or pUL69, pIRS1 and pTRS1 activate early gene expression (Romanowski et al., 1997). In addition, pIRS1 and pTRS1 counteract the interferon response induced by viral infection (Child et al., 2004). Interferon induction in infected cells is designed to shut off protein synthesis, and thereby impede virus replication. Examples of such interferon induced pathways include the protein kinase R (PKR) activation pathway which results in inhibition of translation initiation, and the oligoadenylate synthetase (OAS) pathway which causes degradation of rRNA and tRNA. pIRS1 and pTRS1 block these 2 processes, and ensure viral protein synthesis, although the mechanisms of these actions have yet to be elucidated.

1.4.3.4. Examples of Early (β) gene functions

Most regions of the HCMV genome are expressed during the early phase, and the majority of proteins synthesised with early kinetics function mainly in one of two ways: either by direct involvement in synthesis, cleavage and packaging of the viral DNA and virion assembly, or by providing the correct cellular environment for these events to happen (White and Spector, 2007). Early genes whose products are directly involved in DNA replication include the viral DNA polymerase (pUL54) and its processivity factor (ppUL44), the single stranded DNA binding protein (ppUL57); and the pUL70, pUL102, and pUL105 proteins which form the heterotrimeric helicase primase complex (McMahon and Anders, 2002, Anders and Gibson, 1988, Ertl and Powell, 1992). UL84 is also expressed early in infections, and interacts with IE86 to promote replication from the oriLyt region (He et al., 1992). UL56 and UL89 gene products engage in DNA cleavage (Buerger et al., 2001), and the UL51, UL52, UL77 and UL104 gene products are predicted to play roles in the transport of the empty capsid to the site of encapsidation, and packaging of the newly replicated DNA molecules (White and Spector, 2007).

1.4.3.5. Late gene expression

Late HCMV genes predominantly encode structural proteins, and/or those involved in the assembly and egress of the virus (Demarchi et al., 1980, Wathen and Stinski, 1982). Although strictly defined, late genes are expressed after the onset of DNA replication,

only 2 HCMV genes are known to be expressed with true late kinetics. In contrast, genes encoding most late proteins are active in early infection. The first, pp28, is encoded by UL99, and its expression absolutely requires DNA replication, as a DNA replication-negative mutant failed to produce pp28-encoding mRNA (Depto and Stenberg, 1992). Pp28 is a phosphoprotein component of the tegument, essential for virus replication. Cells infected with a mutant HCMV lacking the UL99 gene produced normal levels of viral DNA and late protein, but tegument-associated capsids failed to acquire envelopes, suggesting pp28 is required for virion envelopment (Silva et al., 2003). The other true-late HCMV gene, UL94, also resides in the tegument of virions, and may have zinc/DNA binding properties (Wing et al., 1996). Some genes, whose products are synthesised late, are also transcriptionally active early on in infection. pUL69 and pUL44 are present in very low amounts at early time points, but are among the most abundant of all infected cell proteins at late points (Geballe et al., 1986). Nevertheless their transcripts are present in equally high concentrations at early and late time points. This suggests post-transcriptional rather than transcriptional regulation is responsible for the late appearance of some late proteins.

1.4.3.6. Assembly and egress

The complex assembly pathway of the HCMV virion takes place both in the nucleus and the cytoplasm, and without arrest of host protein synthesis (Britt, 2007). Final virion assembly takes place in virion assembly complexes (ACs), which are large, highly vacuolated circular structures generally present in one copy per infected cell, and that appear as cytoplasmic inclusions characteristic of HCMV-infected cells (Das et al., 2007). Capsid assembly begins in the cytoplasm, concludes in the nucleus, and requires the presence of the major capsid scaffold protein (pUL80.5) and the minor capsid scaffold protein (pUL80), as well as the final nucleocapsid components (Gibson, 1996). The major capsid protein (MCP, pUL86) lacks a nuclear localization signal (NLS), and its nuclear translocation requires complex formation with pUL80.5 (Plafker and Gibson, 1998). Transport of MCP into the nucleus is followed by further interactions with pUL80 (Plafker and Gibson, 1998). Triplex formation also occurs in the cytoplasm; a homodimer of the minor capsid protein (mCP; UL85) enters the nucleus bound to a single mCP-binding protein (mC-BP; UL46) (Baxter and Gibson, 1997). Extensive intranuclear interactions between MCP, pUL80, pUL80.5, mCP, and mCP-BP then gives

rise to the B-capsid (see 3.2), which becomes the mature C-capsid (nucleocapsid) through DNA encapsidation, and loss of the assembly proteins (Gibson, 1996). Loss of the assembly proteins without DNA packaging leads to the formation of defective A-capsids (Gibson, 1996). Elimination of the scaffold proteins is realised by the proteolytic activity of UL80, which cleaves itself, and pUL80.5. This catalytic activity appears limited to post capsid assembly times, and acts as a trigger for DNA packaging (Gibson, 1996). DNA enters the empty capsid through a portal complex formed by pUL104, which occupies a single vertex in the capsid icosahedron (Dittmer and Bogner, 2005). Unit length DNA molecules are then encapsidated by the cleavage/packaging enzyme complex, which is comprised of UL56 and UL89 (Poon and Roizman, 1993).

DNA-lacking capsids are not detected in the cytoplasm, despite the availability thereof all procapsid constituents. This may be due to insufficient cytoplasmic concentrations of capsid proteins, or the requirement for conformational changes in capsid components brought about during nuclear translocation (Gibson, 1996).

Nuclear egress of the nucleocapsid is not well understood, but the most popular model seems to be that the nucleocapsid undergoes an envelopment step to reach the perinuclear space, followed by a de-envelopment step which releases it into the cytoplasm (Severi et al., 1988). The interaction of two HCMV proteins, pUL50 and pUL53 has been associated with the envelopment and deenvelopment steps (Gibson, 1996).

Perhaps the most poorly defined stage of virion assembly and egress is the incorporation of the tegument. It is thought the nucleocapsid acquires at least part of the tegument in the nucleus, and subsequently loses this tegument layer before nuclear translocation. Strong evidence for partial nuclear tegumentation includes a report of positive pp150 immunostaining at the surface of nucleoplasmic nucleocapsids (Hensel et al., 1995). Further evidence was provided for the nuclear incorporation of pp150, when functional fluorescent tagged-pp150 was shown to colocalise with the MCP at nuclear replication compartments (Sampaio et al., 2005). Despite these findings, the critical role of pp150 is now thought to be in final virion maturation in cytoplasmic AC compartments, and not in initial tegumentation (AuCoin et al., 2006, Tandon and Mocarski, 2008).

Envelopment of the fully tegumented particle is thought to occur through budding into cytoplasmic tubules or vesicles. pUL99 plays an important role in inducing envelopment of particles, as its absence leads to the accumulation of tegumented particles inside the cell (Silva et al., 2003). Envelopment appears to cause a conformational change in the tegument layer, leading to a more solid association of the tegument and envelope proteins, and is followed by subsequent modification steps, such as phosphorylation, glycosylation, and redistribution of envelope components (Gibson, 1996, Nogalski et al., 2007).

1.4.4. Latency and reactivation

Following primary infection, HCMV establishes life-long persistence in the host. HCMV can remain latent in permissive cells, in the absence of detectable production of infectious virus and global viral gene expression. Latency therefore is non-symptomatic, and sporadic reactivation of the virus in healthy individuals generally leads to sedate symptoms, similar to those found with primary infection. Reactivation of the virus can however lead to significant morbidity and mortality in the case of immunocompromised or immunosuppressed individuals, such as AIDS and transplant patients respectively (Sinclair and Sissons, 2006).

Studies on latency and reactivation have identified cells of the myeloid lineage, and specifically monocytes/monocyte-derived cells to be the major sites of latent HCMV, although many other tissues may also harbour latent HCMV genomes (Taylor-Wiedeman et al., 1991). $CD34^+$ stem cells which give rise to the myeloid lineage can also contain viral DNA, but the virus genome has never been detected in non-myeloid $CD34^+$ derived cells. $CD34^+$ progenitors or undifferentiated myeloid cells do not support productive viral infection; studies suggest the HCMV genome resides as a circular episome in these cells, at frequencies only detectable by PCR amplification (about 1 in 10,000 cells) (Bolovan-Fritts et al., 1999, Slobedman and Mocarski, 1999).

The fact that substantial global gene expression has not been observed during latency suggests HCMV can be truly latent, meaning its persistence is not dependent on constant low level productive infection (Sinclair and Sissons, 2006). An obvious and important question therefore is whether HCMV latency is associated with the production of latency-specific transcripts. Addressing this question *in vivo* has been

confounded by the extremely low copy numbers of the virus genome in the blood of healthy seropositive persons, and therefore limited mainly to experimental infection of myeloid lineage cells (Sinclair and Sissons, 2006). Initially, several cytomegalovirus latency-specific transcripts (CLTs) were detected by PCR in the blood of healthy carriers (Sinclair, 2008). These included MIE-expressed transcripts, and a viral homologue of IL-10, transcribed from the UL111A gene (Jenkins et al., 2004). The roles of these CLTs have not been clearly defined. Microarray analysis of cDNA clones isolated from non-permissive $CD34^+$ cells identified a substantial number of genes to be expressed during experimental latency (Goodrum et al., 2002). Many of these transcripts are also expressed during the lytic cycle, and may have represented low level abortive, or undetected productive infection, but the expression of 2 of them was subsequently confirmed in naturally latent cells (Sinclair, 2008). The first, UL138, is detectable in $CD34^+$ undifferentiated cells, and monocytes of healthy seropositives. UL138 mRNA levels decrease upon reactivation of the virus, and its deletion results in restricted latency (Goodrum et al., 2007). Later, an antisense transcript to the UL81 and UL82 genes, named UL81-UL82ast, and its resulting protein product LUNA were detected in monocytes carrying latent virus (Bego et al., 2005), and more recently *in vivo* (Bego et al., 2011). The functions of both UL138 and LUNA remain unknown.

The expression of IE genes seems absolutely required for virus reactivation, as little or no RNA from the IE region can be detected in cells of healthy seropositive individuals (Sinclair and Sissons, 2006). Hence it seems as long as the IE gene expression is sufficiently suppressed, the virus will remain latent, but lytic replication will occur if IE genes are allowed to be expressed. Differentiation of myeloid cells into macrophages and dendritic cells (DCs) is accompanied by reactivation of IE expression (Taylor-Wiedeman et al., 1994). This closely resembles the *in vitro* response of these cells to experimental infection, where differentiation of non-permissive monocytic or $CD34^+$ progenitor cells into mature DCs or macrophages brings about complete permissiveness, and the production of infectious virus progeny. The non-permissiveness of monocytic and $CD34^+$ progenitor cells has been ascribed to the inability of these cells to express viral IE genes, and not to the lack of virus binding and entry (Gonczol et al., 1984). Moreover it has now been accepted that both differentiation-dependent transcription factors, and chromatin structure around the

MIE promoter play a major role in determining IE gene expression (Sinclair, 2010b). An increase in the level of DNAse I hypersensitivity in regions of the MIE promoter, coincided with differentiation and thereby permissiveness of latent cells (Ghazal et al., 1990). A number of cellular transcription factors have been found to upregulate MIE gene expression *in vivo* by binding to the MIE promoter (Sinclair and Sissons, 2006). Also, sequence analysis has predicted the binding of a number of repressive cellular transcription factors to the MIE promoter (Shelbourn et al., 1989). The levels of the repressive transcription factor YY1 drop dramatically following differentiation of cultured non-permissive NT2D1 cells into HCMV-permissive cells, whilst another negative regulator of the MIE promoter, the ERF protein, exerts its influence by recruiting histone deacetylaces (HDACs) to the viral MIE promoter (Sinclair and Sissons, 2006). This repressive effect on chromatin structure can be reversed *in vitro* by the treatment of undifferentiated cells with HDAC inhibitors (Murphy et al., 2002). Such functional remodelling of the chromatin around the viral MIE promoter has been demonstrated *in vivo* in $CD34^+$ cells and monocytes of healthy seropositives, the MIE promoter was associated with methylated histones, but ex vivo differentiation of these cells into mature DCs resulted in demethylation and acetylation of the MIE promoter region (Reeves et al., 2005). This is in accord with *in vitro* chromatin immunoprecipitation assays, which showed that the MIE promoter associates with chromatin silencing proteins in undifferentiated cells, and with histone markers of active transcription upon differentiation and virus reactivation.

The expression of limited HCMV transcripts during latency indicates that the repressive chromatin structure seen around the MIE promoter in undifferentiated $CD34^+$ cells cannot exist for the whole genome (Sinclair, 2010a). In correlation with its activity, and in contrast with the MIE promoter, experimental infection of $CD34^+$ cells revealed an active chromatin structure around the LUNA promoter (Reeves and Sinclair, 2010). Observation of binding sites for GATA-1 and GATA-2, important regulators of gene expression in haematopoetic cells, in the promoter region of LUNA lead to the hypothesis that LUNA expression in latently infected cells may be facilitated by GATA-mediated recruitment of histone acetylases. Indeed post-translational modifications of histones associated with the HCMV genome are now thought to be important not only in latency and reactivation, but also in gene expression during productive infection

(Sinclair, 2010a). Studies of the relationship between chromatin structure and lytic gene expression however, are at an early stage and will likely shed more light on this subject in the near future.

An analysis of the cellular miRNA profile of $CD34^+$ progenitor cells carrying latent virus identified a small number of miRNAs that were significantly downregulated during latency (Poole et al., 2011). Of these, hsa-miR-92a is a known regulator of haematopoiesis and immune function. hsa-miR-92a is known to target GATA-2, which in turn activates cellular IL-10 expression. Consistent with this, GATA-2 and cellular IL-10 levels were found to be upregulated, and it was hypothesised that increased IL-10 levels helped maintain latent carriage of virus. In accord with this, engagement of secreted cellular IL-10 by antibodies caused a fall in the number of latent genomes, possibly due to increased cell death. Cellular miRNA expression may therefore be modulated to optimise the cellular environment for latency, and thereby promote survival of cells harbouring the latent virus.

1.5. Prevalence, transmission, and detection

1.5.1. Prevalence

Since both primary infection and superinfection with HCMV are often asymptomatic, and overt disease is not associated with pathognomonic symptoms, laboratory testing is required to identify carriers. Following primary infection, HCMV-specific antibodies persist in the blood, and their detection provides a simple and efficient method of identifying carriers (Weber et al., 1999, Cannon et al., 2010). Conventional methods of HCMV specific IgM and IgG detection such as the complement fixation (CF) assay, the latex agglutination (LA) assay, and indirect haemagglutination assay (Cannon et al., 2010) have now been replaced by more sensitive enzyme linked immunosorbent assays (ELISAs) (Cannon et al., 2010).

Using such techniques, HCMV seroprevalence is estimated to range between 40% and 100% among different populations worldwide (Ho, 1990). The incidence is higher in

populations living in poor socioeconomic conditions, in developing countries, and in women, who tend to be more readily infected (Ho, 1990). Higher incidence in women could be explained by their increased susceptibility during sexual transmission, or habitual reasons such as higher likelihood of caring for young children, who shed high levels of the virus (Cannon et al., 2010). Race has also been shown to correlate with seroprevalence, with non-whites consistently appearing more likely to be seropositive. It is unclear whether the higher incidence of HCMV in non-whites is solely due to their generally lower socioeconomic status, or if other factors such as cultural or customary practices also play a role (Cannon et al., 2010).

1.5.2. Transmission

Transmission of the virus requires close contact, and can occur horizontally or vertically. Horizontal transfer of the virus generally occurs through infected bodily secretions, sexual contact, and urine (Brown and Abernathy, 1998). Organ transplantations and blood transfusions are also known modes of infection (Brown and Abernathy, 1998). Evidence for the sexual transmission of the virus includes observation of HCMV disease in two men after sexual contact with a woman diagnosed also with HCMV disease (Chretien et al., 1977), higher incidence of the virus in sexually active individuals (Jordan et al., 1973), and the presence of identical strains of the virus in sexual partners (Handsfield et al., 1985). Transmission through close contact and urine is demonstrated by the pronounced rate of seroconversion amongst day care workers (Adler, 1990).

Transplant recipients undergoing immunosuppressive treatment are at a major risk of contracting disease. In seropositive individuals, superinfection and reactivation together occur up to 100% of the time, although reactivation is thought to account for the majority of these cases (Ho, 1990). The frequency of primary infections amongst seronegative organ recipients has been estimated at 53%-73% (Ho, 1990). This naturally follows receipt of an organ containing latent virus, as evidenced by restriction endonuclease analysis (Ho et al., 1975), and the presence of identical strains in the donor and the recipient (Chou, 1986). In one report, multiple HCMV strains were detected in 5 of the 11 renal transplant recipients studied, all of whom were seronegative prior to organ transplantation (Stanton et al., 2005).

Vertical transmission of virus from mother to foetus can occur during primary or recurrent infection of the mother, and mostly involves passage of HCMV virions from maternal blood to the foetus, although entry of infected maternal leukocytes, endometrial, or cervical cells into the foetal circulation, or ingestion by the foetus of infected placental tissue may also lead to congenital infection (Leguizamon and Reece, 1997, Alford et al., 1990). Perinatal infection is more common than congenital infection, and includes passage through an infected uterine cervix during delivery (Ho, 1990) and lactation (Stagno et al., 1980b). Infected children are prolific spreaders of HCMV, as they shed virus for prolonged periods in their urine and other bodily secretions, whereas seropositive adults rarely excrete virus in such high levels and for extended periods (Adler, 1989).

1.5.3. Detection

Transplant and transfusion recipients, newborns and AIDS patients are in major risk of developing HCMV disease (Caliendo et al., 2000). With the availability of better antiviral drugs, rapid and reliable detection of HCMV infection in these individuals has become imperative, as it identifies patients requiring preemptive HCMV treatment (Van der Bij et al., 1988a). Furthermore, timely exclusion of HCMV infection in transplant recipients can minimise overtreatment with immunosuppressive drugs (Van der Bij et al., 1988a). A number of techniques have been described for the detection of HCMV infection. Conventional cell culture techniques involved inoculation of fibroblast cells with patient material, and the observation of characteristic HCMV cytopathic effect (CPE)(Paya et al., 1987). Although these techniques represent a very reliable method of detection, they do not satisfy the need for rapid diagnosis, as definitive exclusion of infection necessitates up to 21 days incubation time (Stagno et al., 1980a, Goldstein et al., 1982). More rapid methods for diagnosis of HCMV disease have therefore been devised. The presence of HCMV particles in blood is an important marker for potential disease, and is strongly associated with disease severity (Van der Bij et al., 1988a). The HCMV early antigen test identifies HCMV particles in blood by immunofluorescence, using anti-IE72 antibody (Schirm et al., 1987). Such recognition of viraemia is performed either directly on blood samples (detection of early antigen fluorescent foci: EAFF) (Schirm et al., 1987), or indirectly in fibroblasts inoculated with purified portions of blood (Stirk and Griffiths, 1987), and can take as little as 24 hours.

Even more rapid and sensitive techniques such as the antigenaemia test, which involves antibody-mediated detection of HCMV antigens in leukocytes within 3 hours, have also been described (Van Der Bij et al., 1988b). Recently several PCR protocols were designed to amplify specific DNA and RNA viral sequences both qualitatively and quantitatively (Caliendo et al., 2000). Quantitative amplification of viral sequences is advantageous as it can be used to monitor response to treatment (Gandhi and Khanna, 2004). Currently, quantitative PCR (qPCR) assays directed against the viral genome and viral mRNA transcripts are the most commonly used methods of identifying patients encountering viraemia and in need of treatment (Caliendo et al., 2000). Despite their rapidity and sensitivity, even the pp65 antigenaemia test and PCR fail to detect HCMV in the blood of some affected patients, such as in 25% of gastrointestinal patients (Boeckh, 2011). For this reason, sampling from other sites such as bronchoalveolar lavage (BAL) fluid, or specific tissue(s) is also used to document disease (Boeckh, 2011).

1.6. Clinical manifestations and treatment

Despite major improvements in recent years in its diagnosis and treatment, HCMV disease continues to cause major morbidity and mortality. In this section a brief account is given of the spectrum of diseases caused by HCMV, and the current treatments available.

1.6.1. Range of diseases

CMV infections rarely manifest themselves as overt disease in the immunocompetent host. However individuals with deficient, suppressed or immature immune defence, such as AIDS patients, transplant recipients and unborn and newborn babies are at severe risk of developing serious, often life threatening disease. The nature of disease is very variable, but commonalities are observed depending on which of the above high-risk groups the patient belongs to. The most typical diseases caused by HCMV are described below.

1.6.1.2. Congenital and perinatal infections

Immaturity of the immune response makes the foetus and the newborn susceptible to HCMV infection (Alford et al., 1990). In fact HCMV is the most common cause of congenital viral infection, seen in about 1% of all live births (Weller, 1971). It is thought between 5% and 10% of congenitally infected children develop symptoms, and of those, 90% suffer from debilitating disease or die soon after birth (Alford et al., 1990). Congenital HCMV disease generally affects multiple organs, and commonly causes hepatomegaly, splenomegaly, microcephaly and mental retardation (Weller et al., 1962). Manifestations in children include loss of sight and/or hearing, CNS damage, tooth defects and growth retardation, and learning disabilities (Alford et al., 1990). Asymptomatic or subclinical congenital infections lead to disease in 5% to 15% of cases, with sensorineural hearing loss, motor defects and mental retardation amongst the most frequent outcomes, usually appearing within 2 years of birth (Alford et al., 1990). Like congenital infections, perinatal infections lead to the shedding of virus for years, but lower amounts of the virus are secreted (Stagno et al., 1983). Newborn children infected with the virus have a far lower risk of developing disease and predominantly remain asymptomatic (Alford et al., 1990).

1.6.1.3. Infections in transplant patients

Transplant patients routinely undergo immunosuppressive treatment to prevent organ rejection. This results in susceptibility to HCMV infection, resulting from reactivation of pre-existing virus in the recipient and/or reactivation of latent HCMV in the transplant organ (Meyers et al., 1986). Patients in greatest risk are seronegatives receiving an organ from a seropositive donor, as they lack pre-existing immunity to HCMV and therefore produce higher viral loads. An exception is bone marrow transplants, where pre-transplant seropositive recipients are in greater risk of developing HCMV infection compared to seronegative patients (Meyers et al., 1986). Seronegative bone marrow recipients meanwhile, are more likely to be infected, when their donors are seropositive. The most common forms of disease in transplant patients include pneumonitis, enteritis and hepatitis (Gandhi and Khanna, 2004).

1.6.1.4. Infections in AIDS patients

Whereas transplant recipients undergo temporary immunosuppression, AIDS patients are terminally immunosuppressed, and are therefore at severe risk of developing HCMV disease (Jacobson and Mills, 1988). HCMV causes a remarkably different range of diseases in AIDS patients compared to other high risk groups. Retinitis is the most common manifestation of HCMV disease in AIDS patients, followed by gastrointestinal disease and encephalitis (Cunha, 2010). Furthermore subsequent progression to AIDS is also accelerated in HIV patients with HCMV disease (Gandhi and Khanna, 2004).

1.6.1.5. Infections in the Immunocompetent host

HCMV is well controlled in the immunocompetent host, such that primary infections in healthy individuals are predominantly asymptomatic. On occasion however primary infection can lead to mononucleosis-like illness (Boeckh and Geballe, 2011), and very rarely to pneumonia or gastrointestinal HCMV disease (Cunha, 2010).

1.6.1.6. Associations with other diseases and cancer

In addition to the typical HCMV diseases discussed in 1.6.1.2-1.6.1.5, HCMV has been linked with a number of other diseases but it is not known whether it is a direct cause of these conditions or simply a bystander (Cunha, 2010). HCMV may exacerbate prognosis of inflammatory bowel disease, new onset diabetes , and atherosclerosis (Cunha, 2010, Lawlor and Moss, 2010). Furthermore it has been suggested that constant surveillance of HCMV in seropositive individuals leads to accelerated immunosenescence, the term given to the weakening of the immune system with age (Pawelec et al., 2010).

The potential link between HCMV and certain forms of cancer has long been subject of debate, and is now gathering evidence (Michaelis et al., 2009). Despite early reports that HCMV can transform human embryonal cells *in vitro* (Geder et al., 1976), occurence of such transformation lacks solid evidence (Michaelis et al., 2009). More conclusive findings have come in recent years, concurrent with the availability of more sensitive methods of detecting HCMV (Michaelis et al., 2009). The presence of viral DNA and antigens, as well as low level persistent virus replication in tumour cells but not adjacent healthy cells was shown in more than 90% of patients suffering from certain types of cancer, such as colon and breast cancer (Söderberg-Nauclér, 2008,

Harkins et al., 2002, Cobbs et al., 2002). In addition, patients whose tumours harboured lower levels of HCMV survived longer than those with high level infection in their tumour cells (Söderberg-Nauclér, 2008). HCMV is generally not regarded as an oncogenic virus, but rather an oncomodulator, meaning it can render tumour cells more malignant (Cinatl Jr et al., 1996). It is thought HCMV can aggravate malignancy by positively contributing to factors such as proliferation, angiogenesis, survival, invasion, resistance to apoptosis, and evasion of immune surveillance (Michaelis et al., 2009). In addition to oncomodulating functions, HCMV has been shown to induce site specific breaks in chromosome 1 at positions 1q21 and 1q42 (Fortunato et al., 2000, Fortunato and Spector, 2003). 1q21 strand breaks are thought to affect an unidentified breast cancer tumour suppressor gene (Bièche et al., 1995) whilst disruption at position 1q42 may ablate the function of ADPRT, a gene involved in DNA repair and replication, whose malfunction correlates with glioblastoma development (Baumgartner et al., 1992, Li et al., 1995). With this in mind, it is hereby suggested that HCMV is predominantly oncomodulating, and mildly oncogenic.

1.6.2. Treatment

Almost all drugs currently used to treat or prevent HCMV disease target the viral replication cycle (Boeckh and Geballe, 2011). Ganciclovir (and its oral alternative valganciclovir) is a guanosine analogue that is phosphorylated by the viral kinase UL97, and causes termination of HCMV DNA replication (Boeckh and Geballe, 2011). Cidofovir and foscarnet are nucleoside monophosphate and pyrophosphate analogues respectively, which also abolish viral DNA replication, but unlike ganciclovir these do not require prephosphorylation (Chou, 1999). Other drugs such as acyclovir have also proved successful in restricting HCMV infection in transplant patients (Boeckh and Geballe, 2011). Unfortunately resistance by HCMV has been observed against all currently available drugs, and is often brought about by mutations in the UL97 kinase gene in the case of ganciclovir, or in the DNA polymerase gene for all anti-DNA replication drugs (Boeckh and Geballe, 2011). Resistance is most common when the levels of the drugs are insufficient to completely suppress HCMV replication. If resistant virus progeny arises some time after commencement of treatment, there will be an initial lowering of viral load followed by an increase in viral load when the resistant form of the virus dominates; or if the virus is resistant to begin with, virus

shedding may persist at original levels several weeks into therapy (Chou, 1999). Moreover, drug resistance is most frequently observed in severely immunocompromised hosts, and those suffering from long term or relapsing HCMV disease. Presence of drug resistance should trigger treatment with alternative drug(s), whilst certain mutations in the DNA polymerase gene can give rise to resistance against multiple drugs (Chou, 1999, Boeckh, 2011).

1.7. Controlling the host

1.7.1. Anti-apoptotic functions

It is well known that apoptosis is used by the host as a method of eliminating virus infected cells (O'Brien, 1998). It is also known that viruses including HCMV have evolved multiple mechanisms to avoid or delay apoptosis of infected cells (Goldmacher, 2005). Apoptosis in response to virus infection can be induced intrinsically, or extrinsically. The intrinsic pathway can be initiated by stress sensors such as p53, causing the release of pro-apoptotic molecules from the mitochondria, and activation of caspases. In the extrinsic pathway, signalling through death receptor molecules by soluble death receptor ligands, and cytotoxic immune cells such as natural killer (NK) cells and cytotoxic T lymphocytes (CTLs) also leads to caspase activation (Goldmacher, 2005). Importantly therefore, intrinsic and extrinsic induction of apoptosis share overlapping downstream features, and apoptosis most often involves cross-talk between, and convergence of, mitochondrial and death receptor signals (Terhune et al., 2007). Caspase activity will then trigger a cascade of reactions, leading to events such as DNA fragmentation, loss of cytoskeletal integrity, and membrane blebbing; resulting in cell death (Terhune et al., 2007).

Due to its relatively long replication cycle, HCMV requires lasting survival of the host cell following infection. HCMV manages to inhibit apoptosis for prolonged periods by encoding an impressive array of anti-apoptotic functions. Interestingly, HCMV anti-apoptotic proteins show no sequence homology, but obvious functional homology to cellular death-suppressor proteins, which suggests such functions have arisen by means other than gene capture (Brune, 2011). In addition, most HCMV anti-apoptotic

functions discovered so far are expressed with IE kinetics, and some are encoded by genes essential for virus replication, highlighting their importance in the life cycle of the virus.

The first anti-apoptotic HCMV proteins to be discovered were IE72 and IE86, each of which is sufficient to block TNF-α induced apoptosis (Zhu et al., 1995, Lukac and Alwine, 1999). However, IE72 and IE86 are not able to protect cells from apoptosis following UV irradiation of the host cell. These molecules are nuclear in localization and are unlikely to directly interact with any cytoplasmic molecules. Furthermore the levels of transcription of many prominent pro-apoptotic and anti-apoptotic cellular genes remained unchanged following expression of IE1 and IE2 genes, leading to suggestions IE72 and IE86 may exert their anti-apoptotic effects through modulation of p53 activity (Zhu et al., 1995). HCMV infection of coronary smooth muscle cells (SMCs) resulted in the upregulation of endogenous p53, which co-localised with IE86 (Speir et al., 1994). Subsequently, the physical interaction between p53 and IE86 was demonstrated by immunoprecipitation. Furthermore IE86 abolished the transactivating effect of p53 on a reporter gene, when all 3 genes were plasmid-cloned and transfected into SMCs. In another study using coronary SMCs, IE86 but not IE72 expressed from an adenovirus vector (Ad-IE-86) inhibited p53-mediated apoptosis when cells were superinfected with Ad-p53, or treated with doxorubicin, a substance known to induce the p-53 dependent apoptotic pathway (Tanaka et al., 1999). In both cases, control cells showed markedly higher rates of apoptosis than IE86-expressing cells. The inhibition of p53 function by IE-86 in SMCs results in continued cell cycle progression, and has been linked with coronary restenosis, a condition in which excessive proliferation of SMCs leads to narrowing of blood vessels and restricted blood flow. Mechanistically, it is known that IE-86 is able to repress p53 function by physically binding p300 and CBP, and blocking their transactivation of p53 (Hsu et al., 2004).

In a screening experiment carried out on HeLa cells, the protein product of exon 1 of the UL37 gene (UL37x1) was identified as an anti-apoptotic gene as it conferred resistance to Fas-mediated, and TNF-α mediated apoptosis (Goldmacher et al., 1999). Due to its mitochondrial localisation, and its physical interaction with the cellular mitochondrial protein adenine nucleotide translocator (ANT), the UL37 protein was

called viral mitochondrion-localised inhibitor of apoptosis (vMIA) (Goldmacher et al., 1999). vMIA displays analogous functions to the cellular anti-apoptotic protein Bcl-2 which also localises to the mitochondria, such as inhibition of cytochrome C leakage from the mitochondria through preserving the integrity of the mitochondrial membrane (Goldmacher et al., 1999).

An early event in extrinsic apoptosis after the engagement of death receptors at the cell surface is the recruitment of procaspase 8, and its subsequent proteolytic cleavage into active caspase 8. Screening of an expression library of HCMV genes lead to the identification of the UL36 gene product as an inhibitor of procaspase-8 activation, and its designation as viral inhibitor of caspase-8 activation (Skaletskaya et al., 2001). vICA represented the first non-essential IE HCMV function, as inactivating mutations in the UL36 gene were reported to be inconsequential to viral replication.

Infection of human fibroblasts with a library of strain AD169 mutants lead to the observation that viruses lacking the UL38 gene underwent excessive cell death, characteristic of apoptosis, and failed to produce maximal virus progeny compared to intact AD169 (Terhune et al., 2004). UL38 over expression protects cells from apoptosis induced by both *dl337*- a mutant adenovirus known to lack a caspase-inhibiting gene that works against intrinsic apoptosis-, and also thapsigargin, which is known to disrupt calcium flux in the ER, but not from apoptosis induced by ligation of the Fas receptor (Terhune et al., 2004). The target(s) of pUL38, or the mechanism by which this molecule performs its death suppressing function remain unknown.

The β2.7 transcript accounts for >20% of virally encoded mRNAs at early times of infection (Section 1.2.3.3), but no protein product has ever been detected for this molecule (Greenaway and Wilkinson, 1987, Spector, 1996). Screening a human cDNA library with β2.7 identified a subunit of the mitochondrial enzyme complex I as a potential interacting partner. Inhibition of complex I activity is known to result in apoptosis, and treatment of neuronal U373 cells with rotenone, a complex I inhibitor, resulted in substantial cell death. Infection with the intact but not the Δβ2.7 Toledo strain of HCMV protected cells from cell death, and β2.7 expression alone was sufficient for this protection. Rotenone treatment of uninfected cells, or cells infected with the mutant virus coincided with the relocalisation of GRIM-19, an essential

subunit of complex I, from a diffuse cytoplasmic location to the perinuclear space. β2.7 was able to abort this relocalisation, and was found to be associated with GRIM-19 by immunoprecipitation. Furthermore this association helped maintain ATP production in the face of oxidative stress during HCMV infection (Reeves et al., 2007).

1.7.1.1. Death receptors and HCMV infection

Since death receptor induced apoptosis is a mode of removing virus-infected cells, it is only logical to predict that HCMV infection may trigger over expression of death receptors at the cell surface, or that HCMV may contain specific functions to inhibit surface expression of pro-apoptotic death receptors.

Tumour necrosis factor alpha (TNF-α) receptor I (TNFRI) is a major mediator of pro-apoptotic and pro-inflammatory signals (Baillie et al., 2003). TNF-α signalling through TNFRI has also shown to induce antiviral and cytotoxic effects in the target cell. It may therefore be beneficial for the virus to block TNF-α signalling by reducing the levels of TNFRI at the cell surface and therefore prevent apoptosis of infected cells (Baillie et al., 2003). Infection of THP1 myelomonocytic cells, and U373 glioblastoma cells with HCMV was shown to significantly downregulate cell surface TNFRI expression (Baillie et al., 2003). This function was abolished when UV-inactivated virus was used, confirming that viral gene expression was required for this effect. It is not known which HCMV gene(s) is responsible for this function.

CD95 (APO-1/Fas) is the best characterised member of the tumour necrosis factor (TNF) superfamily, whose activation by Fas ligand (FasL) can lead to apoptosis (Peter and Krammer, 2003). It has previously been shown that adenoviruses inhibit apoptosis by removal of CD95 molecules from the cell surface (Elsing and Burgert, 1998, Shisler et al., 1997). Surface Fas downregulation by the 10.4K and 14.5K adenovirus proteins has been shown to involve internalization, and degradation in lysosomes (Elsing and Burgert, 1998). CD95 downregulation by the Myxoma virus leukaemia –associated protein (MV-LAP) has also been described, although in the context of viral infection, this effect appears counterbalanced by another Myxoma virus factor, which upregulates CD95 expression (Guerin et al., 2002).

Given the diverse anti-apoptotic functions encoded by HCMV, the question of whether HCMV downregulates CD95 expression seemed a valid one. Previously, it was reported

that infection with the Smith and Eisenhardt clinical strains lead to an upregulation of Fas receptor on the surface of HFF cells (Chaudhuri et al., 1999). The authors also found that mock-infected cells expressed no/undetectable levels of surface Fas. This is unexpected, as abundant Fas expression on the surface of neonatal forskin and adult dermal fibroblasts has been reported (Jelaska and Korn, 1998). I therefore sought to investigate whether HCMV infection would alter surface Fas levels of cultured human fibroblasts.

1.7.3. Immune-evasion

HCMV confronts strong innate and adaptive immune responses throughout its life cycle (Jackson et al., 2011). The importance of these responses is evident from extensive HCMV-related morbidity and mortality in immunosuppressed, immunocompromised, and immunonaïve individuals. At the same time, HCMV employs a variety of mechanisms by which to evade immune surveillance. The efficiency with which HCMV is able to modulate the host's defence mechanisms is exemplified by its life-long persistence following primary infection. Here, a brief overview is given of the immune responses that are elicited through HCMV infection, and the counteracting measures put up by the virus in return, with special emphasis on NK cells.

1.7.3.1. Innate immunity

Innate immunity plays a major role in the fight against invading viruses. The innate immune response detects viral elements by pattern recognition receptors (PRRs), and activates intracellular signalling cascades leading to the secretion of type I interferons and inflammatory cytokines. This activates numerous interferon-responsive and cytokine-responsive genes, and leads to the establishment of an antiviral state to protect neighbouring cells from infection (Takeuchi and Akira, 2009).

Binding and entry of HCMV particles rapidly initiates many signal transduction pathways, leading to a global remodelling of gene expression (Isaacson et al., 2008). HCMV infection can induce cellular gene expression in a number of ways. gB and gH upregulated transcription of Sp1 and NF-κB (Yurochko et al., 1997). Virion proteins such as pp71, together with the IE1 and IE2 encoded proteins have been demonstrated to activate the expression of a number of cellular genes (Liu and Stinski, 1992, Malone

et al., 1990). Differential display analysis identified 15 cellular mRNAs that were upregulated 8 hpi, all of which belonged to interferon-responsive genes (Zhu et al., 1997). Interestingly, these cellular transcripts were also activated by UV-irradiated HCMV, suggesting virion components and not viral gene expression were responsible. DNA microarray analysis comparing transcripts in mock-infected and HCMV-infected cells identified numerous antiviral genes, including interferon-responsive and inflammatory cytokine-responsive genes to be induced by HCMV infection (Browne et al., 2001, Zhu et al., 1998). Moreover, many genes encoding pro-inflammatory cytokines and chemokines are switched on by the NF-κB transcription factor, which is activated following engagement of receptors such as toll-like receptor-2 (TLR2), by viral envelope glycoproteins (Yurochko et al., 1997). Interestingly HCMV has evolved to benefit from these early innate responses too. For example, the presence of NF-κB elements within the MIE promoter indicates NFK-B activation may drive viral gene expression and consequently viral replication (Sambucetti et al., 1989).

1.7.3.1.1. NK-evasion functions

NK cells are a heterogeneous population of cells which express a wide range of both activating and inhibitory surface receptors (Vivier et al., 2011). NK cell cytotoxicity is regulated by the balance between activating and inhibitory signals received from the target cell. NK cells play a vital role in the defence against HCMV infection; this is evidenced by severe HCMV disease in individuals with defective NK-cell functions (Biron et al., 1989). The interaction between HCMV-infected cells and NK-cells has been the subject of intense research, and has revealed much information not only about HCMV infection and pathogenesis, but also about the functioning of NK-cells in general (Wilkinson et al., 2008). A brief account is hereby presented of the array of functions HCMV uses to protect infected cells from NK-lysis.

Natural killer group 2, member D (NKG2D) is a potent activating receptor found on NK-cells and T-cells, which promiscuously binds a number of proteins, such as MICA, MICB, and UL16-binding proteins 1-4 (ULBP1-4) (Champsaur and Lanier, 2010). Virus infection, and specifically HCMV IE1 and IE2 proteins are known to cause activation of NKG2D through upregulation of these ligands (Wilkinson et al., 2008). To defeat this susceptibility, gpUL16 prevents the surface expression of MICB, ULBP1 and ULBP2 by

direct binding and sequestration in the ER to modulate NK-recognition (Dunn et al., 2003a).

An important determining factor in the activation of NK-cells is the level of MHC class I molecules at the surface of potential target cells. According to the 'missing self' hypothesis, constitutive expression of MHC class I molecules on the surface of healthy cells prevents NK-mediated killing. On the other hand, removal of MHC class I molecules from the surface of infected cells serves as a signal for NK attack (Karre et al., 1986). The downregulation of MHC class I molecules on the surface of infected cells by HCMV (see 7.3.2) can therefore leave cells susceptible to NK cytotoxicity (Wilkinson et al., 2008). To counter this, the virus expresses an MHC class I homologue (gpUL18), which like endogenous MHC class I, is able to form trimeric complexes with β2-microglobulin and endogenous peptides (Beck and Barrell, 1988). gpUL18 binds the inhibitory NK receptor LIR1 with 1000 times higher affinity than HLA-I molecules (Cosman et al., 1997, Willcox et al., 2003), but importantly is not subjected to the same downregulation by the US2, US3, US6 and US11 gene products (Park et al., 2002). Inhibition of $LIR1^{+}$ NK cells by gpUL18 has been observed with multiple donors (Prod'homme et al., 2007).

The inhibitory surface receptor CD94/NKG2A is expressed on the surface of most human NK cells, where it interacts specifically with the non-classical MHC class I molecule, HLA-E (Braud et al., 1998). Surface expression of HLA-E in turn requires prior binding with a peptide derived from residues 3 to 11 of the signal sequence of MHC class I molecules (Braud et al., 1997). This binding occurs in the ER, following cytoplasmic cleavage and processing of classical MHC class I molecules, and the TAP-dependent transportation of the signal sequence to the ER. Thus HLA-E recognition by NK cells monitors normal MHC Class I synthesis and TAP function. Contrary to expectation, TAP-inhibition by US6 (see 1.7.3.2) did not result in suppression of HLA-E surface expression during productive HCMV infection (Tomasec et al., 2000). This is because the UL40 signal peptide (SP UL40) of HCMV carries a nine amino acid sequence, exactly homologous to endogenous HLA-E binding peptides, which strongly promotes HLA-E surface expression in a TAP-independent manner (Tomasec et al., 2000). Moreover, an HCMV UL40 deletion mutant underwent increased NK cytolysis

compared to virus expressing a functional UL40 protein (Wang et al., 2002). UL40 therefore acts as an important NK-evasion function.

As described in 1.7.3.1, the interferon pathway is important in the innate defence of cells against HCMV infection. Interferons secreted from infected cells bind their cognate receptors on neighbouring cells, leading to the activation of interferon sensitive genes (ISGs). The product of the UL83 gene, pp65, has been shown to block the induction of many ISGs and pro-inflammatory genes very early after infection, and virus lacking the UL83 gene induced a stronger expression of many interferon responsive genes (Browne and Shenk, 2003). To achieve this, pp65 has been proposed to inhibit the activation of NF-κB and IRF1, but the mechanism of this function remains to be elucidated (Browne and Shenk, 2003). In addition to suppressing antiviral gene expression, pp65 has been shown to directly bind the activating NK receptor NKp30, but remarkably, such binding did not result in NK lysis, but rather diminished the potential of NK cells to kill virus-infected cells (Arnon et al., 2005). Since pp65 is neither secreted from infected cells nor present on the surface of infected cells, or the viron envelope, direct NK binding may require its release by lysis of infected cells (Wilkinson et al., 2008). This suggests NK-lysis of some infected cells in an infected tissue may benefit the overall survival of the virus, by decapacitating roaming NK cells. The highly passaged AD169 and Towne laboratory strains have undergone 15-kb and 13-kb deletions corresponding to the UL/*b*' region (Cha et al., 1996). Although these strains protected cells from NK-killing, this protection occurred to a lesser extent than seen with clinical isolates (Cerboni et al., 2000, Wang et al., 2002). This prompted the prediction that the UL/*b*' region encoded one or more NK-evasion functions that had been lost from certain strains through extensive culture *in vitro* (Tomasec et al., 2005). This prediction was validated when a recombinant Towne virus in which the Toledo strain UL/*b*' had been inserted (Towne/Tol11) displayed protection from NK-killing in levels comparable with wildtype virus (Tomasec et al., 2005). Ectopic expression of individual UL/*b*' genes in fibroblasts then lead to the identification of UL141 as an NK-evasion gene, as UL141-expressing cells exhibited enhanced resistance to NK-mediated cytolysis (Tomasec et al., 2005). The observation that HCMV strains containing UL141, but not those lacking the UL141 gene downregulated CD155 expression, and that gpUL141 alone could downregulate CD155 demonstrated that CD155 was targeted by

gpUL141 (Tomasec et al., 2005). CD155 is a cell surface ligand for the NK-activating receptors CD226 and CD96 (Tomasec et al., 2005). It was further determined that gpUL141 blocked the maturation of CD155 from an Endo H-sensitive intracellular form to an Endo H-resistant cell surface species (Tomasec et al., 2005).

In addition to binding CD155, CD226 also recognised CD112 expressed on the surface of infected cells (Bottino et al., 2003). It was recently shown that CD112 levels are also downregulated during the course of HCMV infection, and that UL141 was essential but not sufficient for CD112 degradation, suggesting other viral factors are involved (Prod'homme et al., 2010).

Sequence homology between the UL142 gene, also in the UL/*b*' region and UL18, and the *in silico* prediction that UL142 encodes an MHC class I-related molecule prompted investigations into the potential role of UL142 as an NK-evasion gene (Wills et al., 2005a). Cells expressing UL142 consistently inhibited NK-killing, and furthermore, small interfering RNA (siRNA) knockdown of UL142 lead to increased NK-mediated lysis of cells in the context of HCMV infection (Wills et al., 2005a). In a later study UL142 was found to downregulate the full length MICA allele, but not its truncated MICA *008 version from the surface of infected cells, and thus bestowed NK-resistance to infected cells (Chalupny et al., 2006). It has subsequently been clarified that downregulation of MICA from the cell surface by UL142 is a result of sequestration in the *cis*-Golgi (Ashiru et al., 2009).

14 HCMV microRNAs (miRs) have so far been discovered (Dhuruvasan et al., 2011). Prediction of the target sequences of HCMV miRNAs using an algorithm lead to the identification of MICB as a potential target for miR-UL112 (Stern-Ginossar et al., 2007a). MICB is a stress-induced ligand for the activating receptor NKG2D, and is critical for the killing of virus infected cells by NK cells (Stern-Ginossar et al., 2007a). Ectopic expression of miR112 resulted in MICB downregulation in various human cell lines (Stern-Ginossar et al., 2007a). Furthermore, MICB downregulation resulted in reduced binding of NKG2D to target cells, and more importantly, in reduced NK-killing (Stern-Ginossar et al., 2007a). Inhibition of NK-killing by miR112 revealed a new mode of immune-evasion by HCMV. In addition to being very target-specific and non-

immunogenic, miRs may be simpler to develop from an evolutionary point of view, than functional proteins (Stern-Ginossar et al., 2007a).

That HCMV devotes a large number of its gene products to escape from recognition by NK cells is an indicator of the significance of NK-cells in innate immunity against HCMV infection. Nevertheless, it is important to recognise, that many of the functions classified above as NK-evasive, are also relevant with respect to other immune effector cells (Wilkinson et al., 2008). This is because NK-receptors such as NKG2D and LIR-1 are also found on other cell populations including certain subsets of T cells (Wilkinson et al., 2008).

1.7.3.2. Adaptive immunity

The adaptive immune response can contribute substantially to the control of HCMV disease, as evidenced by the therapeutic benefit of administering immunoglobulin or adoptive transfer of cytotoxic T lymphocytes (Khan et al., 2002).

Conflicting reports have been produced on the role of antibodies in the immune response against HCMV. Pre-existing maternal antibodies play an important role in preventing congenital infection of the foetus (Fowler et al., 1992). Also, infants born to seropositive mothers were more resistant to transfusion borne HCMV infection (Yeager et al., 1981). In one study, intravenous injection of anti-HCMV antibodies to renal transplant recipients reduced the incidence of primary HCMV disease from 60% to 21% (Snydman et al., 1987). It is unclear whether antibodies protect allogeneic transplant patients from HCMV disease (Jackson et al., 2011). What is clear though is that primary infection with HCMV ensues production of antibodies specific to several HCMV proteins. These include the pp65 and pp150 tegument proteins, the IE1 protein, gB and gH envelope glycoproteins, and also different components of the gH/gL/UL128-UL131A pentameric envelope complex (Rasmussen et al., 1991, Wills et al., 1996, Macagno et al., 2010, Schoppel et al., 1998).

Primary infection with HCMV elicits extremely strong and broad $CD4^+$ and $CD8^+$ T cell responses, which forever leaves a fingerprint in the T cell compartment (van de Berg et al., 2008, Sylwester et al., 2005). These changes include much higher frequencies of circulating cytotoxic T cells (van de Berg et al., 2008). $CD8^+$ and $CD4^+$ T cells are an integral part of the immune response to HCMV (Miller et al., 2001). Administration of

ex vivo expanded $CD8^+$ T cells restored immunity to HCMV in immune-deficient individuals (Riddell et al., 1992), whilst the reconstitution of $CD8^+$ T cells in transplant patients inversely correlated with HCMV viraemia (Barron et al., 2009, Jackson et al., 2011). Moreover, specific $CD8^+$, and $CD4^+$ T cells have been detected to the majority of HCMV ORFs, with the strongest responses occurring against pp65, IE1, pp28, pp150, gB and gH (Jackson et al., 2011). Evidence exists for the early and lasting accumulation of pp65 specific $CD8^+$ T cells following HCMV infection (van de Berg et al., 2008). It is estimated that HCMV specific $CD4^+$ and $CD8^+$ T cells comprise up to 10% of their respective populations in the peripheral blood (Sylwester et al., 2005). HCMV specific $CD4^+$ T cells appear within a week of detection of viral DNA in blood, and produce T helper type 1 (Th1) cytokines, augment $CD8^+$ T cell, and B cell responses, and may also contribute to direct killing of infected cells (Jackson et al., 2011, Miller et al., 2001). Also, the majority of $CD8^+$ T cells displayed high levels of *ex vivo* cytotoxicity against their respective targets (Wills et al., 1999).

HCMV subverts detection by $CD4^+$ and $CD8^+$ T cells by a number of mechanisms. The downregulation of MHC class I molecules at the surface of HCMV infected cells begins at IE times (Yamashita et al., 1993), and leads to complete annihilation of IE antigen presentation, and a great reduction in pp65 presentation to $CD8^+$ T cells, (Besold et al., 2007). For this function, HCMV uses a family of homologous genes in a sequential multistep process, each of which is alone able to lessen surface expression of MHC class I molecules (Ahn et al., 1996). In noninfected cells, MHC class I heavy chains form trimeric complexes with β2-microglobulin and antigenic peptides, which are transported to the cell surface, in a transporter associated with antigen processing (TAP) dependent manner (Ahn et al., 1996). In HCMV-infected cells, US3 retains class I heavy chain molecules in the ER, US6 prevents peptide loading by inhibiting TAP-mediated transport of peptides into the ER, and US2 and US11 proteins cause the specific export of class I heavy chains into the cytoplasm to be degraded by proteasomes (Ahn et al., 1996, Wiertz et al., 1996b, Wiertz et al., 1996a, Jones et al., 1996). Down-modulation of class I molecules by US2, US3, US6 and US11 proteins represents an elegant immune subversion function. Although these proteins are each independently capable of causing reduced presentation of antigens, they utilise overlapping rather than redundant means of achieving this function, and their

coincidence enables the virus to more efficiently downregulate class I surface expression (Jackson et al., 2011).

$CD4^+$ T cells too are targets of HCMV evasion. This includes the inhibition of inducible and constitutive MHC class II expression at the transcriptional, post transcriptional, and post-translational level (Miller et al., 2001). Furthermore, it was recently recognised that during latency, the viral homologue of the IL10 cytokine (CMV IL10), is able to downregulate MHC I and II expression, and obstruct cytokine production by infected cells; thus tergiversating the $CD4^+$ T cell response (Jackson et al., 2011).

To summarise, HCMV is a master regulator of the cellular environment, able to expertly subjugate the broad, persistent, and powerful response it receives from the immune system in admirably artistic ways. In the presence of a healthy immune system, these subjugations are sufficient only for the maintenance of the virus in a latent state, but lapses in the body's defence against HCMV quickly and commonly lead to productive HCMV infection and disease.

1.8. Aims

The HCMV genome encodes at least 170 protein-coding genes (Gatherer et al., 2011), many of which remain functionally uncharacterised. Studies aiming to discover HCMV gene function are confounded by this large number of genes and their complex expression pattern. In addition, though the genes responsible for numerous immune-modulatory functions have already been discovered, many others await identification. In the past, the discovery of most immune-evasive HCMV genes has involved algorithmic predictions or fortuitous experimenting (Wilkinson et al., 2008). Some HCMV genes carry out essential core functions, and therefore it is not possible to generate mutants of every HCMV gene. Also, due to the fact that HCMV grows poorly *in vitro*, HCMV research is largely limited to fibroblasts, and even in these cells preparation of high titer virus stocks remains labour intensive. There exists a compelling need for a high-throughput system with which to map HCMV functions to individual genes. It was hypothesised that if all HCMV genes were individually cloned into an efficient mammalian expression vector, they could be rapidly screened for a specific immunodulatory action in a gain-of-function assay. The expression vector

selected for gain of function screening was the AdZ (adenovirus with zero cloning steps) vector generated by Dr Richard Stanton (Stanton et al., 2008b). A second approach also available was to use a panel of deletion mutants, each lacking a block of non-essential genes, in loss of function assays. The aims of the project are outlined below:

1- Assist in the generation of the AdZ library encoding all HCMV protein coding genes
2- Characterise and validate expression from this library
3- Investigate results of particular interest
4- Characterise the HCMV gene function responsible for Fas downregulation

2-Materials and Methods

2.1. Cells

Human foetal foreskin fibroblasts (HFFFs) selected for efficient growth of HCMV were kindly provided by Dr Graham Farrar (Porton Down). HFFFs were immortalised by transduction with the human telomerase reverse transcriptase (hTERT), as described by McSharry et al (McSharry et al., 2001), by Sian Llewelyn-Lacey and designated HFFF-hTERTs. To facilitate efficient infection with adenovirus vectors, HFFF-hTERTs were engineered to express the human Coxsackie adenovirus receptor (hCAR) (Leon et al., 1998, Stanton et al., 2008a). 293 cells (Graham et al., 1977) expressing the tetracycline repressor protein (293-TREx)were purchased from Invitrogen (Invitrogen, R71007).

2.2. Cell culture

Cells were grown in Dulbecco's modified Eagle's medium (DMEM, Invitrogen, 41965) and kept in humidified incubators at 37^{o}C at 5% CO_2 concentration. Media was supplemented with 2% penicillin streptomycin (Invitrogen, 15070) and 9% FCS (Invitrogen, 10500), referred to in the text as complete medium, unless stated otherwise.

Mycoplasma screening was undertaken by Sian Llewelyn-Lacey and using the VenorGeM® Mycoplasma PCR detection kit (Biochrom AG, Germany). Any mycoplasma positive cell lines or virus stocks provided to the laboratory were cleaned or discarded, and only mycoplasma negative stocks were used.

Cells were routinely maintained in 150 cm^2 flasks (Fisher, TKV-123-051F), from here on referred to as T150 flasks, or Cell bind T150 flasks (Fisher, TKV-123-050G), and allowed to reach confluency. To split adherent cells, media was removed, the monolayer rinsed in 10 ml of PBS (Fisher, 14190), and 5 ml of trypsin (Fisher, 25300) was added until cells detached (~3 minutes). Cells were then resuspended by the addition of 5 ml of complete medium, and the resulting suspension of cells was normally divided by a factor of 4 for HFFF cells, and 10 for 293 cells, before being seeded in fresh flasks.

2.2.1 Cell counting

To estimate the number of cells in a suspension, the sample was mixed, and 10 µl were immediately removed and used to fill the counting chamber (Fisher, MNK-790-M). Cells within a grid representing 0.1 µl were counted under x10 magnification, and the average number obtained. Cell numbers per ml of the suspension were calculated by multiplication by 10^4. The required volume of cells was then transferred to the appropriate flasks, topped up with more media if necessary, and allowed to adhere overnight.

2.3. Molecular biological techniques

2.3.1. Polymerase Chain Reaction (PCR)

PCR was carried out to amplify predefined regions of DNA using specific primers purchased from Sigma and Invitrogen, and using the Biometra T3000 Thermal cycler. PCR programs were adjusted according to the length of DNA to be amplified, the amount of DNA required, and the type of polymerase applied. Concentration of template DNA varied depending on preparation used. Buffers, primer, DNA and DMSO were defrosted at 37^{o}C in the water bath, while dNTPs were defrosted at room temperature. All tubes were then kept on ice until returned to -20^{o}C. Sample PCR reaction mixtures and programs used are listed in appendix I.

2.3.2. Primer and oligonucleotide design

Primers were designed to have a melting temperature of at least 60^{o}C. Palindromic, and dimer-forming sequences were avoided as much as possible, except where positions were fixed, such as reverse primers for amplifying Merlin genes, in order to clone into the AdZ vector. The sequences of primers and oligonucleotides used are listed in appendix II.

2.3.3. DNA gel electrophoresis

PCR products were resolved based on size on tris-acetate-EDTA (TAE) gels (Fisher, BPE1332-20) containing 0.8-1% agarose (AGTC Bioproducts, A4-0700) and visualised by ethidium bromide staining (Sigma, E1510-10 ml) under UV light. Agarose was dissolved

in 50 ml of 1x TAE buffer, 2.5 µl ethidium bromide were added in the fume hood (final concentration 0.5 µg/ml), and the solution poured into a gel tray taped at either end, and containing a comb, and allowed 30 minutes to solidify. The comb and the tapes were removed and the tray was transferred into a gel tank containing 1xTAE buffer to above the level of the gel. DNA loading dye (Therm-oscientific, R0611) was added at 1x working concentration to the DNA samples, which were pipetted into the wells alongside 10 µl of molecular weight markers (Eurogentec, MW-1700-10). Electrophoresis was undertaken for 1-12 hours and at 10-100 V depending on the resolution required.

2.3.4. Gel purification of DNA

Under UV light (Spectroline transilluminator, model TVC-312A), appropriate size DNA bands were excised, transferred to sterile 0.5 ml Eppendorf tubes, and their weights determined. DNA was then purified from the gel using the illustra GFX PCR DNA and Gel Band Purification kit (28-9034-70, GE Healthcare). Here, 1 µl capture buffer was added for every milligram of the gel, and the gel melted at 60°C. The solution was transferred to a column, and placed in a fresh Eppendorf tube. Once 30s elapsed, samples were spun at 13K rpm for 30s (Biofuge Fresco, Heraeus). The flow-through was discarded and 500 µl of the wash buffer was pipetted onto the column. A further 30s spin at 13K rpm (Biofuge Fresco, Heraeus) was performed, and the column was transferred into a fresh Eppendorf tube. 30 µl of elution buffer were pipetted onto the column, and following a further 30s incubation, DNA was eluted by a 1 minute spin at 13K rpm (Biofuge Fresco, Heraeus).

2.3.5. Determination of DNA concentration

DNA concentrations were determined using the NanoDrop ND-1000 spectrophotometer. 1 µl of water was loaded onto the measurement pedestal to initialize the nanodrop. Then 1 µl of the solvent was loaded to establish a blank measurement, and finally 1 µl of the sample was loaded onto the nanodrop and DNA concentration was obtained in ng/ ul.

2.3.6. Propagating SW102 Escherichia coli

SW102 Escherichia coli contain lambda phage Red recombination genes, which are heat activated to mediate recombination between homologous stretches of DNA. Propagation at 37^{o}C therefore causes inopportune recombination within the *E.coli* genome, culminating in the death of the bacterial culture. For this reason, SW102 *E.coli* were grown at 32^{o}C. SW102 *E.coli* contain a chloramphenicol resistance gene, and chloramphenicol was always included at 12.5 ug/ml concentration. Ampicillin was included in the culture at 50 ug/ml, except where the *ampicillin resistance/Lac*Z*/Sac*B cassette had been removed.

LB broth was prepared by dissolving 20g of LB powder (Melford, Melford laboratories Ltd, L1703) per 1L of distilled water. Solutions were autoclaved and allowed to cool to at least 32^{o}C. Bacteria from a frozen glycerol stock were inoculated into appropriate volume of LB, or when being grown using an existing culture, 5 µl were inoculated into fresh LB and the culture was grown in a shaking incubator at 32^{o}C.

2.3.7. Generating glycerol stocks

Bacteria were grown overnight in 5 ml LB containing antibiotic(s). 1 ml aliquots of the overnight culture were transferred to fresh 2 ml pre-assembled white-cap tubes (Greiner Bio One, 723 261) in triplicates, and vortexed with 0.5 ml glycerol until mixed homogeneously. Glycerol stocks were then frozen at -80^{o}C. Glycerol stocks were generated for all adenovirus (Ad) BAC vectors, and HCMV-BAC recombinants. This is important as it provides a reproducible source of the recombinant vector.

2.3.8. Pouring plates

LB broth was prepared by adding in 1g of tryptone (Fisher, BPE1421-500), 0.5g of yeast extract (Sigma, Y4000), and 1.5 g of agar (Fisher, S/0370/63) per 100 ml of distilled water. The solution was autoclaved and allowed to cool to about 60^{o}C. The appropriate antibiotics were then added (Chloramphenicol was always included at 12.5 ug/ ml of the solution), and the solution was poured into Petri dishes (20 ml each, Fisher, PDS-140-050F). The agar was allowed to solidify for 20 minutes at room temperature. Plates were then inverted and allowed to dry at 32^{o}C for a further 20 minutes, and were subsequently kept at 4^{o}C until used.

2.3.8.1. Positive and negative selection of the *ampicillin resistance*/*Lac*Z/*Sac*B cassette

The cassette of selectable markers in SW102 *E.coli* contains a *Sac*B gene, an *ampicillin resistance* gene, and a *Lac*Z gene. *Sac*B encodes levansucrase, which in the presence of sucrose efficiently inhibits *E.coli* growth. *Lac*Z gives rise to blue colonies when X-gal and IPTG are available in the media.

When negatively selecting against these markers, 5% sucrose (Fisher, BP2201) was dissolved in the LB broth solution prior to autoclaving. 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-gal, (80 ug/ml final concentration) and Isopropyl β-D-1-thiogalactopyranoside (IPTG, 200 nM concentration) were also added prior to the pouring of plates. This ensured that bacteria containing the selection cassette, but mutated in *Sac*B gene were identified by the production of blue colour.

For positive selection of the above markers, ampicillin was incorporated at 50 ug/ml, together with X-gal (80 ug/ml final concentration) and IPTG (200 nM final concentration). Here, surviving blue colonies indicated the loss of the selection cassette. Figure 2.1 shows an example of a positive selection plate containing blue and white colonies.

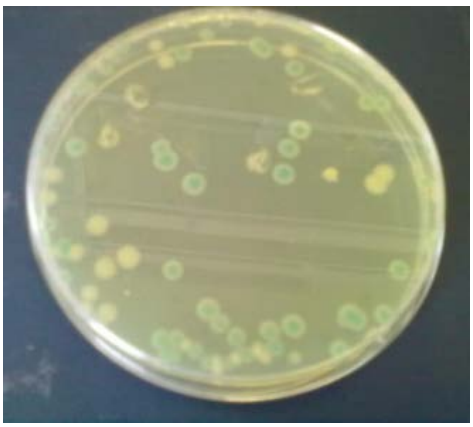


Figure 2.1

Screening for bacteria containing the selection cassette

Blue and white colonies represent bacteria with and without the selection cassette respectively. In positive selection blue colonies are picked, whilst white colonies are used in negative selection

2.4. Purification of BAC DNA from SW102 cultures

2.4.1. Minipreparation of BAC DNA

The AdZ bacterial artificial chromosome (BAC) and the Merlin BAC carried by the SW102 *E.coli* were the targets of gene cloning and manipulation by recombineering. In order to identify positive recombinants and verify the BAC sequence following recombineering, BAC DNA was purified from 5 ml cultures of SW102 *E.coli* (Miniprep) using the Qiagen 27106 miniprep kit. Colonies were inoculated and grown overnight in 5 ml of LB containing chloramphenicol (12.5 ug/ml), and where appropriate ampicillin (50 ug/ml). Bacteria were centrifuged at 4000 rpm for 5 minutes (Megafuge 1.0 R, Heraeus instruments). Supernatants were discarded and the cells resuspended in 250 µl buffer P1 (27106, Qiagen) , then lysed and their DNA denatured by the addition of 250 µl S2 buffer (27106, Qiagen) for 4 minutes, before the reaction was neutralised by 250 µl S3 buffer (27106, Qiagen). Samples were mixed by inversion and centrifuged at 13K rpm (Biofuge Fresco, Heraeus) for 10 minutes at room temperature (Biofuge Fresco, Heraeus). Supernatants were transferred to fresh tubes, and DNA precipitated with 0.75 ml isopropanol at 4^{o}C by centrifugation for 10 minutes at 13krpm (Biofuge Fresco, Heraeus). Pellets were washed in 500 µl 70% ethanol, with recovery ensured by centrifugation for 5 minutes at 13 krpm (Biofuge Fresco, Heraeus), and the pellet dried at 37°C in a heating block. DNA was then resuspended in 30 µl of elution buffer (EB) (27106, Qiagen).

2.4.2. Maxiprepreparation of BAC DNA

In order to obtain transfection quality DNA, recombinant DNA was purified using the Nucleobond BAC 100 kit (NZ740579, Fisher Scientific). 250 ml of culture was grown overnight, transferred to 250 ml polycarbonate centrifuge bottles (Fisher CFS-300-520C), and centrifuged at 6000 rpm for 10 minutes at 4^{o}C (SLA-1500 rotor, Sorvall Evolution RC). Supernatants were discarded and pellets were resuspended in 12 ml of S1 buffer. Cells were then lysed with 12 ml S2 buffer for 4 minutes with gentle swirling of the tube. 12 ml N3 buffer was then used to neutralise the lysis reaction. The

resulting mixture was cooled on ice for 5 minutes and centrifuged at 6000 rpm for 10 minutes at 4^{o}C (SLA-1500 rotor, Sorvall Evolution RC). This allowed separation of solid cell debris from released cell contents. The supernatants were passed through a filter-paper, and the flow-through put through a column, which was previously equilibrated with 6 ml N2 buffer. The column was then washed twice in 18 ml N3 wash buffer, and the DNA eluted with 15 ml N5 buffer, pre-heated to 50°C. The elution was collected in 30 ml polypropylene tubes (Thermo Scientific, 03719), mixed with 11 ml room temperature isopropanol, and centrifuged at 15000 rpm and 4°C for 30 minutes (SS-34 rotor, Sorvall Evolution RC). Supernatant was discarded and pellet washed in 5 ml 70% ethanol, and centrifuged at 15000 rpm, at 15 minutes (SS-34 rotor, Sorvall Evolution RC). Ethanol was allowed to evaporate from the pellet for 15 minutes, and DNA resuspended in 100 µl elution buffer. The eluate was then transferred to a fresh tube, and DNA concentration was measured.

2.5. Restriction digest

*Bam*HI restriction digestion analysis was perfomed on miniprepared BAC DNA to verify the integrity of the vector, and also the presence/size of genes cloned into the AdZ vector. Predicted molecular weights were obtained for *Bam*HI-digested DNA fragments by identifying *Bam*HI sites in the DNA sequence (Vector NTI, Invitrogen).

8 µl purified DNA were incubated with 1 µl buffer E (Promega, R005A) and 1 µl *Bam*HI enzyme (Promega, R6021), and incubated for 2 hours at 37°C. Samples were then loaded onto an agarose gel, and electrophoresed at 150V for 90 minutes.

2.6. DNA sequencing

Sequencing primers were designed to bind about 100 bp upstream of either end of the region to be validated. Sequencing was carried out using the *BigDye® Terminator v3.1* Cycle Sequencing Kit (Applied Biosystems, 4337456). A 10 µl reaction mix was prepared containing 4 of BigDye mix, 5 µl purified BAC DNA, and 1 µl primer (3.2 pmol/µl). The following thermocycling program was used:

95°C for 5 minutes, then 100 cycles of 95°C for 30seconds, 55°C for 10 seconds, 60°C for 4 minutes.

Completed reactions were cleaned using Performa DTR Gel Filtration Cartridges (EdgeBio, 90780) according to the manufacturer's instructions. Briefly, cartridges were centrifuged for 3 minutes at 3.1 krpm (Biofuge Fresco, Heraeus) transferred to fresh tubes, and samples were added to the column. Centrifugation was performed for 3 minutes at 3.1 krpm (Biofuge Fresco, Heraeus) and the eluate was retained for sequencing. Sequencing was carried out by Central Biotechnology Services (CBS, Cardiff University), and analysed using the Vector NTI software (Invitrogen).

2.7. Recombineering

Recombineering (recombination-mediated genetic engineering) makes use of lambda phage red recombination genes in SW102 *E.coli*, which upon heat-activation; promote recombination between DNA sequences with as little as 30bp of homology (heat-activation occurs at between 37°C, and 45°C, and is optimal at 42°C). This allows one-step cloning using purified PCR products or oligonucleotides, and enables insertion, deletion, or replacement of pre-defined sequences. Figure 2.2 contains a diagrammatic representation of these recombineering events.

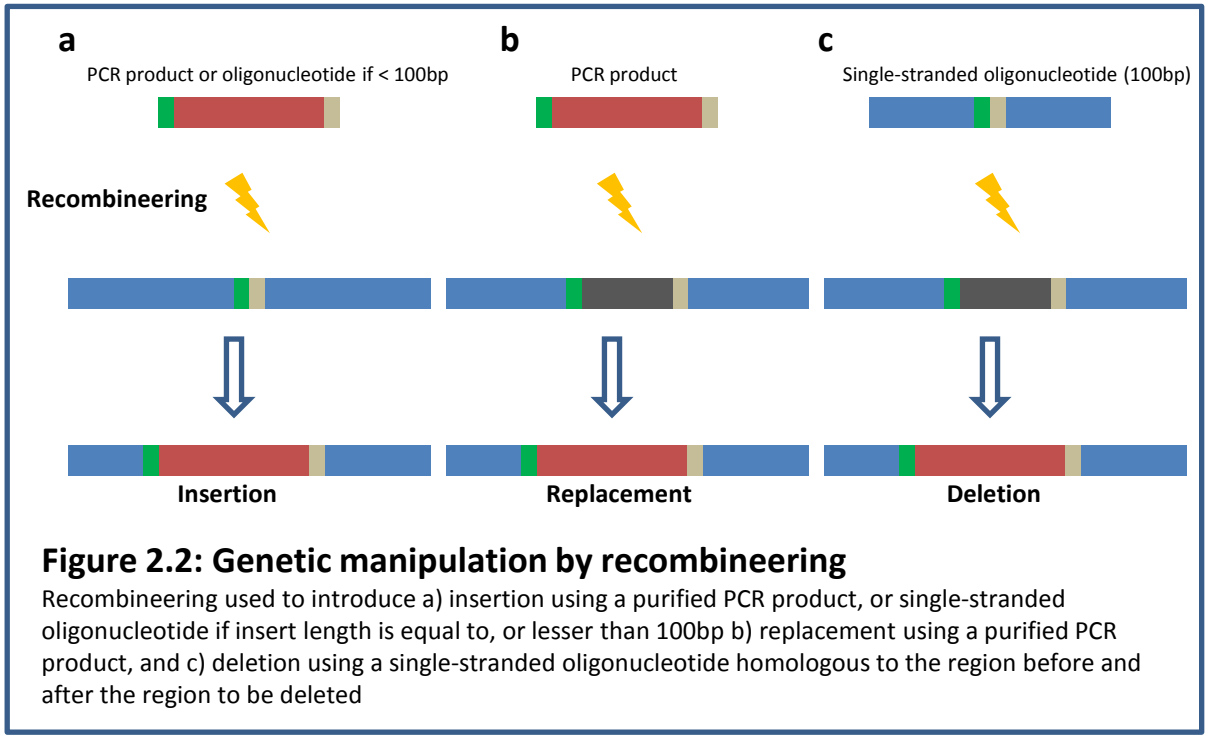


Figure 2.2: Genetic manipulation by recombineering

Recombineering used to introduce a) insertion using a purified PCR product, or single-stranded oligonucleotide if insert length is equal to, or lesser than 100bp b) replacement using a purified PCR product, and c) deletion using a single-stranded oligonucleotide homologous to the region before and after the region to be deleted

SW102 *E.coli* containing the vector of interest were grown overnight in 5 ml LB containing appropriate antibiotic(s). 0.5 ml of the overnight culture was transferred to 25 ml LB, containing chloramphenicol, and where appropriate ampicillin, and allowed to reach the optical density (OD) of 0.6 at 600nm (about 3 h, measured using the Ultraspec 3000, Pharmacia Biotech spectrophotomer) in a shaking incubator at 32^{o}C. The culture was transferred to a 50 ml falcon (Fisher, CFT-900-031F), and the lambda phage genes were activated through a 15 minute heat shock in the 42°C water bath. The culture was inverted every 5 minutes to ensure good heat-mixing. Bacteria were then cooled by rocking on ice for 15 minutes, and centrifuged at 4000 rpm for 5 minutes at 0°C (Megafuge 1.0 R, Heraeus instruments). Supernatants were discarded, and pellets resuspended in 1 ml sterile ice-cold water by gentle shaking. 25 ml of ice-cold water were added and bacteria were centrifuged at 4000 rpm for 5 minutes at 0°C (Megafuge 1.0 R, Heraeus instruments). This wash step was repeated once more, supernatant was discarded, and the pellet containing competent bacteria was resuspended in 400 µl ice cold water. 25 µl of this resuspension was transferred into a fresh tube, and mixed with DNA containing homology to the target sequence in the vector (4 µl purified PCR product, or 1 µl oligonucleotide), transferred to a pre-cooled 0.2 cm gap electroporation cuvettes (Bio-Rad, 165-2082), and left on ice for 5 minutes. Electroporation was then carried out using the Biorad Micropulser (program EC2).

For negative selection (Section 2.3.8.1), electroporated bacteria were transferred to a fresh universal tube, diluted in 5 ml LB and following a 3 hour recovery, 100 µl of the suspension were plated on selective media. This recovery time is important as it allows the loss of SacB gene products from bacteria that have undergone deletion of the selection cassette, thus preventing their killing in the presence of sucrose. For positive selection using the *ampicillin resistance/Lac*Z*/Sac*B cassette, electroporated bacteria were diluted in 1 ml LB and recovered for 1 hour before being plated on selective media.

2.8. Procedures using Adenovirus

2.8.1. Transfection of 293-TREx cells with adenovirus BAC DNA

Recombinant Ad-vectors were replication deficient due to the deletion of the essential E1 region. In order to generate adenovirus stocks, Ad-vectors were transfected into 293-TREx cells expressing E1 functions. 293-TREx cells also express the tet-repressor gene, which binds to tet-operons in the vector, preventing transgene expression. This provides protection against toxic gene products, or those interfering with Ad-replication.

Transfection was carried out using Polyfect (301105, Qiagen), or Effectene (301425, Qiagen) transfection kits. Plaques were visible 7-10 days after transfection and cells were collected in 2 ml media by scraping, about 12 days post transfection, when appearing fully infected. Optimally infected cells were bright, fully rounded-up, losing contact with neighbouring cells, and easily detached from flask. Flasks were agitated until cells were in a suspension.

2.8.1.1. Polyfect:

$2x10^6$ 293-TREx cells were seeded in a cellbind T25 plate (Fisher, TKV-123-031L), and allowed to adhere overnight. 4 µg maxiprepared DNA were diluted in 100 µl DMEM. 40 µl of polyfect reagent was then added and the solution incubated for 10 minutes at room temperature. Cells were supplemented with 3 ml fresh media, and the DNA complex was added to cells in 1 ml of fresh media. Media were changed 24 hour post transfection.

2.8.1.2. Effectene

$2x10^6$ 293-TREx cells were seeded in a cellbind T25 (Fisher, TKV-123-031L) plate and allowed to adhere overnight. On day of transfection, 1 µg of DNA was diluted with the DNA condensation buffer to a volume of 150 µl. 8 µl of enhancer solution were added and the mixture was vortexed for 5 seconds. Following a 5 minute incubation at room temperature, 25 µl of Effectene transfection reagent were added, and samples were incubated for 10 minutes at room temperature to allow complex formation. Growth

medium was aspirated from the cells, and replaced with fresh medium. 1 ml of fresh medium was then used to transfer the transfection complex onto cells.

2.8.2. Expanding adenovirus stocks

To obtain high titer stocks of virus for experimental use, 293 TREx cells were grown to near-confluency in 5x T150 Cellbind flasks. Cells in each flask were infected with 200 µl of the initial virus stock (2.8.1). When the monolayer appeared fully infected (4-10 days), cells were detached by agitation or using cell scrapers (Greiner Bio one, 541870). Cells were centrifuged at 1500 rpm for 5 minutes (Megafuge 1.0 R, Heraeus instruments), and pellets were combined by resuspension in 20 ml PBS. A further 5 minute centrifugation was performed at 1500 rpm (Megafuge 1.0 R, Heraeus instruments), supernatants were discarded, and the resulting pellet stored at -80°C until purified (see Section 2.8.3).

2.8.3. Purification of adenovirus on Cesium-Chloride (CsCl) gradients

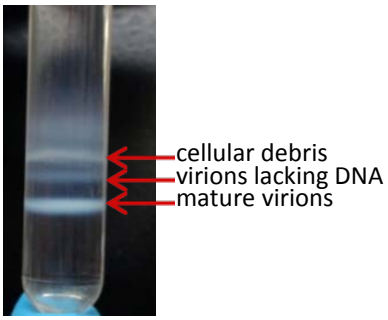
In order to eliminate cellular debris, crude virus (see Section 2.8.2) was centrifuged through a discontinuous Cesium-Chloride (CsCl, Sigma, C3011-500G) gradient; pure virus formed a band at the gradient junction, and was extracted.

1.6 ml heavy solution (3.6M: 1.45g/ml CsCl in 5mM Tris HCL, pH7.8) were pipetted into a 14x 89 mm ultra clear Beckman centrifuge tube (331372, Beckman Coulter), and overlaid with 3 ml of the light solution (1.6M: 1.33g/ml CsCl in 1mM EDTA, 5mM Tris HCL, pH7.8) drop by drop, to cause minimal disturbance in the bottom layer.

To release virus from cells, the cell pellet was resuspended in 7 ml PBS. 7 ml tetrachloroethylene was added, and the mixture shaken vigorously for 30 seconds to form a single phase, followed by centrifugation at 2200 rpm for 20 minutes (Megafuge 1.0 R, Heraeus instruments). The tetrachloroethylene descended to the bottom of the tube, and was separated by a layer of cellular debris from the top PBS layer which contained the virus. The upper aqueous layer containting virus particles was collected, and pipetted onto the Cesium-Chloride gradient. Tubes were then centrifuged at 23000 rpm for 2 hours (SW41 Ti rotor, Ultra Beckman L8-M ultracentrifuge). Following centrifugation, the virus was observed as an opalescent band resting at the interface

between the 2 CsCl solutions (Figure 2.3). The virus band was then extracted using a 19G gauge needle inserted in the side of the tube and transferred to a fresh tube.

Figure 2.3: Purification of adenovirus by CsCl gradient centrifugation
Following centrifugation, adenovirus virions were visualised as an opalescent band residing between heavy and light CsCl solutions. Defective virions formed a band of lower density than virions, but higher density than cellular debris.



2.8.4. Removal of CsCl from virus stocks by dialysis

To avoid the toxic effects of cesium on infected cells, CsCl purified virus (see Section 2.8.3) was dialysed against a buffer solution before being used experimentally. Pores in the dialysis tubing allowed free flow of CsCl out of the tube until an equilibrium was reached with the dialysis buffer. Sequential dialysis of virus preparation in 1L of dialysis buffer (twice) ensured sufficient purification of virus from CsCl.

The CsCl purified virus was diluted to a total volume of 2 ml in sterile dialysis buffer (1mM MgCl2, 135mM NaCl, 10mM Tris HCl, 10% glycerol, pH7.8). Dialysis tubing (Medicell International Ltd, DTV12000.01.000) was pre-warmed by submerging in a beaker of lukewarm distilled water, and was allowed to cool to room temperature. The tubing was then immersed in dialysis buffer and closed at one end by a clip and a knot. Virus was pipetted into the tube through the open end, and was subsequently trapped by applying a clip and a knot to the opening of the tube. The tube was then placed in 1 L of dialysis buffer and positioned on a rocker. The dialysis buffer was removed after 2-5 hours, and replaced with fresh buffer for an overnight incubation at 4°C on rocker. Dialysed virus was removed from the tube by cutting below the clip and transferred into a fresh tube. Virus was then diluted to 10 ml with sterile dialysis buffer, and 1 ml aliquots were prepared in pre-assembled white-cap tubes (Greiner Bio One, 723 261), and stored at -80°C.

2.8.5. Adenovirus titration

2.8.5.1. Spot assay

Once recombinant adenovirus (RAd) stocks were generated, it was necessary to determine the virus titers. Titration of purified adenovirus stocks was based on the spot assay (Bewig and Schmidt, 2000), in which virus particles are quantified immunocytochemically, using a primary antibody directed against virus hexon proteins, and a horseradish peroxidase labelled secondary antibody. In the presence of 3, 3'-diaminobenzidine (DAB) substrate, infected cells were stained brown. The spot assay has a number of advantages over other common methods of adenovirus titration. Compared to the plaque assay which can take up to 10 days, titers can be obtained 48 hours after infection. Furthermore titrating adenoviruses by the spot

assay is an easier procedure than the fluorescent focus assay, which requires fluorescent microscopy. In this method, 293 cells were seeded in 12 well plates (Fisher, TKT-520-070H) at $5x10^5$ cells per well in 1 ml complete media, and allowed to adhere overnight. Virus was serially diluted to obtain 10^{-4} and 10^{-5} dilutions. First, 10 µl of the purified virus stock were diluted in 990 µl of media. Then, 10 µl of the resulting 10^{-2} dilution were diluted in 990 µl of media to obtain 10^{-4} dilution of the virus. 100 µl of this solution were diluted in 900 µl of media to achieve the 10^{-5} dilution. 100 µl of each dilution were added to wells in duplicates. Plates were gently swirled to ensure equal distribution of virus, and incubated at 37°C. 48 hours after infection, culture media were aspirated, and the cells were allowed to dry in a laminar flow hood for 15 minutes, followed by a 15 minute fixation in ice cold 1:1 methanol and acetone at -20°C. The fixation and all subsequent washes were carried out in 1 ml per well, whilst antibody incubations were performed in 0.5 ml for 1h at 37°C and under constant rocking. The fixing reagents were aspirated, and cells washed 3 times in PBS containing 1% BSA (Sigma, A77906-100G). Polyclonal goat anti-adenovirus antibody (Chemicon International, AB1056) was added at 1/5000 dilution, cells were washed 3 times in PBS/BSA, and incubated with donkey anti-goat HRP (Santa Cruz Biotechnology, SC-2056), 1/1000 dilution. Three more washes were performed, and Infected cells detected by the ImmunoPure metal enhanced DAB Substrate kit (34065, Thermo Scientific). A 1x working solution of the DAB/Metal Concentrate was prepared by diluting the concentrate in the stable peroxide solution provided. 0.5 ml of the resulting mixture was added to each well, and incubated at room temperature for 10 minutes. The substrate was then aspirated and replaced with 1 ml PBS. Infected cells were stained brown, and were counted under the microscope. Titres were calculated using the below formula:

Titer (pfu/ml) =average number of infected cells per field of view x number of fields per well /virus volume used (ml) x dilution factor

2.8.5.2. Picogreen fluorescence

Unlike the spot assay, which measures the number of infectious particles, RAd titration by picogreen fluorescence was used to determine the total concentration of virus particles. Picogreen is a fluorescent dye, which displays enhanced fluorescence upon binding to double stranded DNA. Therefore the level of fluorescence can be used to

determine the amount of virus DNA present, and thereby estimate virus titres (Murakami and McCaman, 1999). This method utilises the stoichiometric relationship that exists between virus particle number and virus DNA amount. Importantly, using this method, titres are calculated with reference to a standard curve, generated by applying picogreen to a serial dilutions of a virus of known titre.

2.8.5.2.1. Generating the standard curve:

To generate the standard curve, a series of 2-fold dilutions of the reference RAd was prepared in TE buffer (10mM Tris / 1mM EDTA, pH 8) to cover the range 10^7-10^9 pfu/ml. This represented the working range of the assay and the plate reader. In a 96 well plate (Fisher, TKT-180-070U), 45 µl of each dilution were mixed with 5 µl of lysis buffer (0.5% SDS in TE buffer), and incubated for 15 minutes at room temperature. 20 µl of each sample were then added to 180 µl of picogreen working solution (1/360 dilution of picogreen stock (Invitrogen, P7581) in TE buffer) in duplicates, in black MaxiSorp 96 well plates (Thermo Scientific, 456537). Once mixed, samples were incubated in the dark for 5 minutes for binding to occur, and absorbance was measured at ex485/em538 in the FLUOstar Omega microplate reader (BMG LABTECH). The plate reader was adjusted to use auto-gain, and to measure absorbance by transmitting 10 flashes from the top of the plate. Lysis and picogreen binding of test viruses were achieved as for the reference virus, using 1/10 and 1/100 dilutions. Titres were calculated by comparing absorbance against the standard curve.

2.9. Procedures using HCMV

2.9.1. Generation of recombinant HCMV from HCMV BAC DNA

Purified HCMV recombinant BACs were transfected into HFFFs to produce new virus progeny. Transfection was achieved using the Amaxa™ Basic Nucleofector™ kit for fibroblasts (VPL-1002, Lonza) according to the manufacturer's instructions. Briefly, 10^6 HFFFs were centrifuged at 600 rpm for 10 minutes. Supernatant was discarded and pellet resuspended in 100 µl of Nucleofector solution. 2 µg of purified HCMV BAC DNA were added and the mixture transferred to a cuvette (provided with kit). Cells were

electroporated in the Nucleofector II machine (Amaxa Biosystems, program T16). Samples were then recovered in 5 ml media and transferred to a T25 flask. Plaques appeared about 3 weeks post transfection. Once all cells were detached or appeared dead, media containing virus was collected (about 20 ml) and stored at -80°C until used to expand virus stock (see Section 2.9.2).

2.9.2. Preparation of HCMV stocks

When growing HCMV from a previous stock of known titre, HFFFs were allowed to reach confluency and then infected at MOI ≤ 0.01. Cells were grown in 1 or more T150 flasks, or in Cell Factories (TKT-175-020M, Fisher scientific), depending on the growth properties of the virus strain and the amount required. When growing virus from an initial transfection-produced stock of unknown titre (see Section 2.9.1), 1 ml was used to infect cells in each T150 flask, and 10 ml were used to infect cells in a cell factory. Once cytopathic effect (CPE) was observed for about 80% cells, media were collected, and centrifuged in 250 ml polycarbonate centrifuge bottles (Fisher, CFS-300-520C) at 14000 rpm and room temperature. Media was collected every 24-48 hours, until complete destruction of the monolayer was observed. Following centrifugation, supernatants were discarded and pellets were dislodged with a bacterial inoculation loop, and collected in 1 ml fresh media. The remainder of the pellet was then collected in a further 1 ml fresh media and added to the collected pellet. To resuspend virus present in the pellet, the suspension was homogenised slowly 5-10 times by passing through a 21g needle until loss of all visible clumps. Samples were then centrifuged at 1000 rpm for 1 minute (Megafuge 1.0 R, Heraeus instruments) to pellet the cellular debris, and supernatants were transferred to a fresh tube. Virus was then stored at -80°C until all collections were performed, at which time frozen samples were defrosted at 37°C, pooled together, and aliquoted in pre-assembled white-cap tubes (Greiner Bio One, 723 261).

2.9.3. Titration of HCMV stocks by plaque assay

HFFFs were seeded into 6 well plates (Fisher, TKT-520-030T) at a density of $4x10^5$ per well, and allowed to adhere overnight. The next day, 2 ml each of 10-6, 10-7 and 10-8 dilutions of the virus were prepared, and cells were infected in duplicates of 1 ml with each dilution. Virus from a stock of known titer was included as a control for the

experiment. Infection was carried out under constant rocking for 2 hours at room temperature. Virus was then aspirated and cells in each well overlaid with 10 ml of a 1:1 mixture of 2% avicel (FMC Biopolymer) and 2X media (see below). Plates were then incubated for 2 weeks. The density of the overlay restricted free transfer of released virus from infected cells through the media and limited infectivity to cell to cell contact. This gave rise to plaques, each one resulting from a single infectious HCMV particle. Two weeks after infection the overlay was aspirated and washed away from the monolayer with PBS. Plaques were counted under the microscope and averages between duplicates were obtained.

Titre (pfu/ml) = average number of plaques per well / dilution factor (e.g. 10^6) x volume of virus used for infection (ml)

2% Avicel-type RC591NF stock: 20g of Avicel were resuspended in 1L of water and autoclaved.

2x medium (Per 100 ml): 50 ml sterile water, 20 ml 10x DMEM (Gibco 21430), 20 ml FCS (Invitrogen, 10500), 6 ml sodium bircarbonate (Gibco 25080), 4 ml Pen/Strep (Gibco15070063), 2 ml glutamine (Gibco 25030024)

1% overlay was obtained by mixing equal amounts of the 2% Avicel stock solution and the 2x medium.

2.9.4. Purification of HCMV virions on NaT gradient

To purify HCMV virions from cellular debris and other contaminants, virus was centrifuged through a continuous sodium-tartrate (NaT) gradient, forming a band at its buoyant density, which was subsequently extracted.

Na-phosphate buffer: Two 0.04 M solutions of sodium-dihydrogenphosphate (solution A) and disodium-hydrogenphosphate (solution B) were prepared in water. 19 ml of solution A were mixed with 81 ml of solution B to obtain 100 ml of phosphate buffer with pH 7.4.

Two solutions of differing Na-tartrate solution were then made using the Na-phosphate buffer:

35% Na-tartrate: 35 g Na-tartrate, 65 g Na-phosphate buffer

15% Na-tartrate: 15 g Na-tartrate, 30 g glycerol, 55 g Na-phosphate buffer

The linear gradient was created by a gradient-maker (Hoefer SG50), which allowed mixing of the light and heavy solutions in different ratios. First, with both stopcocks closed, 5 ml of the light solution was pipetted into the reservoir chamber. The connector stopcock was briefly opened to allow the light solution to travel to the edge of the mixing chamber, and closed again. Then, 4 ml of the heavy solution was added to the mixing chamber. Magnetic stir bars were placed in both chambers and the gradient-maker was placed atop a magnetic stirrer. Stirring in the mixing chamber ensured instant mixing of the incoming light solution with the heavy solution, whilst that in the reservoir chamber prevented backflow from the mixing chamber. Both stopcocks were opened simultaneously and solution was let out of the gradient-maker. The initial concentration of liquid exiting the gradient-maker was equal to the heavy solution, and the final concentration equal to that of the light solution. To ensure that solutions were delivered uniformly and consistently, a peristaltic pump (P-1, Pharmacia Fine Chemicals) was used. Solution pumped out of the gradient-maker was used to fill an SW41 ultra-clear tube. 0.5 ml virus (Section 2.9.2) was then overlaid atop the gradient drop by drop, and tubes were filled to 5 mm of the top with sterile PBS. Samples were centrifuged at 23000 rpm for 45 minutes (Ultra Beckman L8-M ultracentrifuge), and virion bands were extracted using 21 g needles (figure 2.4).

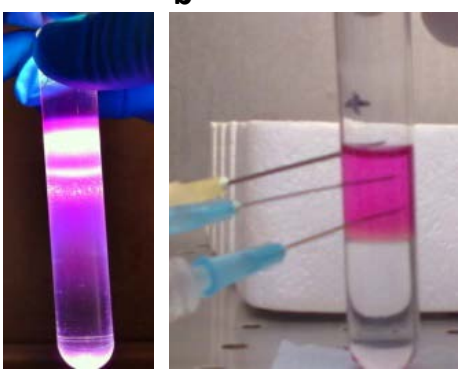


Figure 2.4: Visualisation and extraction of HCMV viral particles following purification on a Na-tartrate gradient

a). NIEPs, virions and dense bodies (DBs) were separated and are observed as distinct bands. NIEPs have the lowest density and form the top band, whilst dense bodies form the lowest band. Virions have intermediate density, and form the middle band. b) Band extraction using needles.

2.10. Infection of cells with Adenovirus and HCMV

The day before infection, cells were counted and seeded, and allowed to adhere overnight. When infecting with Adenovirus, the virus was removed from -80^{o}C, and immediately transferred to a 37^{o}C waterbath. Once defrosted, virus was kept on ice, mixed by gentle pipetting, and the desired volume diluted in complete media. The volume was determined by the following formula:

Volume (ml) = multiplicity of infection x number of cells / virus titer (pfu/ml)

Media was removed from cells and replaced with media containing virus. Infections were carried out for 2 hours at 37^{o}C under continuous rocking, after which virus was removed from cells and replaced with complete media. Virus was then returned to -80^{o}C.

HCMV infections were carried out identically to above, except once defrosted, virus was kept at room temperature.

2.11. Analysis of protein expression

2.11.1. Immunofluorescence

Fibroblasts were seeded in 96 well glass bottom plates (Scientific Laboratory Supplies, TIS5524) at 5000 per well. The following day, cells were infected with HCMV or adenovirus at the desired MOI(s). 72 hours after infection, media was removed and cells were then fixed in 4% paraformaldehyde for 10 minutes, with 2 washes in 100 µl of PBS either side of fixation. Permeabilisation and blocking were carried out in 100 µl of wash buffer for 1 hour with a change of buffer every 20 minutes. All antibody incubations and washes were carried out in the wash buffer unless stated otherwise. Primary antibody incubation was performed for 1 hour at room temperature or overnight at 4°C in 100 µl. Cells were then washed 5 times and the secondary Alexa Fluor-594 goat anti-mouse IgG antibody (Invitrogen, A11020) was applied at 1/500 dilution for 30 minutes at room temperature in a total volume of 100 µl. DAPI (Sigma-

Aldrich, D-9564), when used, was added together with the secondary antibody at 400 ng/ml final concentration to visualise the nuclei. Phalloidin stain conjugated to green-fluorescent Alexa Fluor 488 dye (Invitrogen, A12379) strongly stains F-actin filaments and was added together with the secondary antibody to visualise actin filaments and the outline of the cell (60 units per 96 well plate). Excess secondary antibody, DAPI and Phalloidin stain were removed by 5 washes in the wash buffer. Cells were then washed 5 times in PBS and fixed in 2% paraformaldehyde for 10 minutes. Fixing solution was replaced with 1% paraformaldehyde and fluorescence was detected under a Leica fluorescent microscope (Leica, Germany) at 40x magnification with oil. Fluorescent pictures were obtained using a Hamamatsu camera (Hamamatsu City, Japan), and images acquired with Improvision Openlab 3 software (Improvision, Coventry).

When capturing images, exposure was adjusted until background level fluorescence was minimal or not visible.

4% Paraformaldehyde: 200 ml of PBS were heated to 70^{o}C. In the fume hood, 8 g of paraformaldehyde powder (Sigma, P6148) were added to the hot PBS solution. The powder was mixed into PBS by swirling, and any remaining residues were dissolved by the addition of NaOH (Fisher, S/4800/60). The solution was then placed in ice water, and once at about room temperature, the pH was adjusted to 7.0-7.4. The solution was then filtered and stored at 4°C.

IF buffer: PBS, containing 1% BSA, 0.2% Saponin, and 0.05% sodium azide (saponin was substituted with 0.2% Triton X-100 when permeabilising cells infected with RAd-UL122 and UL123).

2.11.2. Western Blot

Fibroblasts were seeded at 8x10^{5} per T25 flask, or 4x10^{5} per well in a 6 well plate well plates and infected the following day at the desired MOI(s), in 1 ml or 0.5 ml complete media respectively, for 2 hours on an oscillating platform. Virus was removed, and cells were rinsed and incubated in 5 ml DMEM. At 72 hours p.i., cells were rinsed, scraped off (Greiner Bio one, 541870) in cold PBS, and centrifuged at 1500 rpm (Megafuge 1.0 R, Heraeus instruments) for 3 minutes. Cell pellets were resuspended in 200 µl of 1x NuPAGE LDS sample buffer -made from 4x concentrate (Invitrogen, NP0008) in water-

and 50mM DTT. Sonication was carried out for 20s at 20% power (Vibra-Cell, VCX130, Sonics and Materials) to shear DNA, and samples were transferred to PCR tubes and boiled for 5 minutes. Samples were then centrifuged at 13000 rpm for 10 minutes (Biofuge Fresco, Heraeus). Supernatant samples, when analysed were boiled and centrifuged in 1x NuPAGE LDS sample buffer and 50mM DTT, but were not sonicated. As a standard, 20 µl of the final sample were loaded onto each well of a 10% NuPAGE® Novex Bis-Tris gel (Invitrogen, WG1201BOX), and electrophoresed alongside 10 µl of prestained protein markers (Invitrogen, LC5800) for 2 hours at 120V. Some proteins with small predicted sizes were electrophoresed on 4-12% NuPAGE® Novex Bis-Tris gels (NPO321BOX, Invitrogen). Western transfer was carried out for 2 hours at 10 Volts in the Trans-Blot SD semidry transfer cell (Bio-Rad) using 2x transfer buffer made from 20x solution (NuPAGE® Transfer Buffer) in water, membranes were rinsed in distilled water, incubated in antibody extender (Thermo Scientific, 32110) for 10 minutes, rinsed again 5 times in water, and blocked overnight at 4°C in TBST 5% milk. Membranes were then washed 5 times in TBST for 5 minutes each time, and probed with primary antibody overnight at 4°C and in block solution. Primary antibodies used for western blotting are listed in table 2.1. Unbound primary antibody was removed by 5 washes in block solution, each one 5 minutes long, and then horse radish peroxidase -conjugated secondary antibody was applied for 1 hour at room temperature (goat anti-mouse HRP: Bio-Rad, 170-6516 or goat anti-rabbit HRP: Bio-Rad, 170-6515). To eliminate excess secondary antibody, 5 repeated TBST washes of the membrane were completed. Membranes were rocked in 8 ml of SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, 34078) for 5 minutes at room temperature. Illuminated signals were visualised in the GelDoc-It transilluminator (UVP), or exposed to an Amersham Hyperfilm MP (GE healthcare, 28906843) and developed by the compact X4 processor (Xograph Imaging systems)

Transfer buffer (per 500 ml): 50 ml 20x NuPAGE® Transfer Buffer, 50 ml methanol, 400 ml deionised water

TBST (per 1 L): 20 ml 1M TRIS (pH 7.8), 29 g NaCl, 0.5 ml TritonX-100 (Fisher, BP151-500), 0.5 ml Tween 20 (Fisher, BP337500) in water

Block solution: TBST containing 5% dry milk powder

Table 2.1: Primary antibodies used in immunofluorescence and Western Blot

Antigen	Company, Cat number	Dilution	Secondary antibody
Immunofluorescence			
V5 tag	AbD Serotec, MCA1360	1/1000	1/500 Alexa Fluor-594 anti-mouse IgG
Strep tag	IBA-lifesciences, 21507001	1/1000	1/500 Alexa Fluor-594 anti-mouse IgG
IE1, IE2	Pierce-antibodies, MA1-7596	1/1000	1/500 Alexa Fluor-594 anti-mouse IgG
UL14	mAb generated in our laboratory	1/5000	1/500 Alexa Fluor-594 anti-mouse IgG
Western blot			
V5 tag	AbD Serotec, MCA1360	1/2000	1/2000 anti-mouse HRP
human CD155	A.Nomoto, 5D1	1/10000	1/25000 anti-mouse HRP
MHC class-I	HC10	1/2000	1/5000 anti-mouse HRP
human Fas	Santa Cruz, Sc-715 Abcam, ERP5700	1/2000 1/1500	1/2000 anti-rabbit HRP 1/2000 anti-rabbit HRP
UL141	mAb generated in our laboratory,	1/10000	1/10000 antimouse HRP
pp65 (UL83)	Abcam, ab49214-1	1/50000	1/50000 anti-mouse HRP
gB (UL55)	Abcam, ab49214-1	1/25000	1/25000 anti-mouse HRP
Actin	Sigma, A-2066	1/10000	1/10000 anti-rabbit HRP
GAPDH	Abcam, 9485	1/50000	1/50000 anti-rabbit HRP
Strep tag	IBA-lifesciences, 21507001	1/1000	1/2000 anti-mouse HRP
IE1, IE2	Pierce-antibodies, MA1-82057	1/10000	1/2000 anti-mouse HRP
UL14	mAb generated in our laboratory	1/20000	1/2000 anti-mouse HRP

2.11.3. Deglycosylations

Where deglycosylation was to be performed, western blot samples were processed as above, except they were sonicated, boiled, and centrifuged in 1x glycoprotein denaturing buffer (made from 10x concentrate in water: New England Biolabs, B1704S). Digestions were performed overnight at 37°C with EndoH (New England Biolabs , P0702L) or PNGaseF (New England Biolabs, P0704L) enzymes. G5 buffer (New England Biolabs, B1702S) was included in the EndoH reactions, and G7 (New England Biolabs, B3704S) and 10% NP40 (New England Biolabs, B2704S) buffers were added to the mixture in PNGase digestions, according to manufactuer's instructions. Enzymes and buffers amounts used are shown in table 2.2.

Following digestion, 14 µl of each sample were added to 5 µl of 4x concentrate LDS sample buffer (Invitrogen, NP0008), and 1 µl of 1M DTT (50mM final DTT concentration), boiled for 5 minutes, and centrifuged for 2 minutes at 13 krpm (Biofuge Fresco, Heraeus). Western blot was performed as described in section 2.11.2.

Table 2.2: **Deglycosylation reaction mixtures**

	Mock-digested	EndoH-digested	PNGaseF-digested
	Volume (µl)	Volume (µl)	Volume (µl)
Sample	11	11	11
EndoH	-	1	-
PNGaseF	-	-	1
G5 buffer	1.5	1.5	-
G7 buffer	-	-	1.5
NP40	-	-	1.5
Water	2.5	1.5	-
Total volume	15	15	15

2.11.4. Coomasie staining

Protein samples were prepared, processed and electrophoresed as described for western blot (see Section 2.11.2). Following electrophoresis, the gel was incubated in fixing solution for 10 minutes on rocker, and at room temperature. The fixing solution was then replaced with the staining solution without stainer B for 10 minutes. Stainer B was added, and incubation continued for 3-12 hours. Once staining was complete, gel was rinsed and kept in deionised water for a minimum of 7 hours, and for up to 3 days. Images were then obtained by scanning the gels.

Fixing solution (per 100 ml):

40 ml deionised water, 50 ml methanol, 10 ml acetic acid.

Staining solution (per 100 ml):

55 ml deionised water, 20 ml methanol, 20 ml stainer A (Invitrogen 46-7015), 5 ml stainer B (Invitrogen 46-7016)

2.11.5. Silver staining

Protein samples were prepared, processed and electrophoresed as described for western blot (see Section 2.11.2). The polyacrylamide gel containing protein samples of interest was washed in deionised water for 5 minutes, with gentle shaking at room temperature. This wash step was repeated, and water was decanted. Samples were then fixed in fixing solution for two 15 minutes periods, with a change in solution in between. Gel was washed twice with the ethanol wash solution, for 5 minutes each time. Two 5 minutes washes were performed in deionised water, and the gel was incubated in sensitizer working solution – prepared by mixing 1 part SilverSnap sensitizer (Thermo-scientific, 1859065), with 500 parts deionised water (50 µl sensitizer with 25 ml water) for exactly 1 minute. Gel was then rinsed twice in water, and incubated for 5 minutes in a 1:100 mix of SilverSnap enhancer (Thermo-scientific, 1856220), and SilverSnap stain (Thermo-scientific, 1856218) (0.25 ml of enhancer and 25 ml stain), followed by 2, 1 minutes washes in water. Developer working solution was prepared by mixing 0.25 ml of enhancer and 25 ml of developer solution (Thermo-scientific, 1856219), and added to the gel until protein bands of the desired intensity

were obtained (normally 1-3 minutes). To stop further staining, the stop solution was added, and the gel was scanned to attain image.

Fixing solution (per 100 ml): 60 ml water, 30 ml ethanol, 10 ml acetic acid

Stop solution (per 100 ml): 95 ml water, 5 ml acetic acid

Ethanol wash (per 100 ml): 90 ml water, 10 ml ethanol

2.12. Flow cytometry

Cells were seeded and/or infected in T25 flasks or 6 well plates. On the day of staining, media were aspirated, and cells were rinsed in PBS. Cells were then detached in 1 ml trypsin, which was deactivated by the addition of 5 ml complete media. The suspension was transferred to a fresh 15 ml tube, and centrifuged at 1200 rpm for 3 minutes (Megafuge 1.0 R, Heraeus instruments). Supernatants were discarded, and the pellet resuspended in 100-300 µl of cold wash buffer (PBS/1% BSA/0.05% sodium azide). Resuspended cells were transferred to a 96 well plate (Fisher, TKT-180-050D) and centrifuged at 1500 rpm for 3 minutes (Megafuge 1.0 R, Heraeus instruments), and at 4°C. From this point on, all washes and incubations were carried out on ice, and using ice cold solutions. Centrifugation steps were all performed at 4°C, 1500 rpm (Megafuge 1.0 R, Heraeus instruments), and for 3 minutes. Supernatants were poured off and cell pellets resuspended in wash buffer containing the primary antibody. Incubation with the primary antibody lasted 20-30 minutes, and at the midway point, the suspension was remixed by pipetting. Samples were then centrifuged, and supernatants discarded. To remove excess primary antibody, cells were washed twice by repeated resuspension in fresh wash buffer, centrifugation, and removal of the supernatant. Incubation with the secondary antibody was identical to that with the primary, but samples were protected from light throughout. Samples were then washed 3 times in wash buffer, 3 times in PBS, and fixed in 2% paraformaldehyde for 10 minutes. The FACSCalibur (Dakocytomation, Cambridge) or the BD Accuri^{TM} C6 flow cytometers was used to detect fluorescence, and results were analysed using the cell quest (BDIS, San Jose, USA) and Cflow softwares respectively.

Primary antibodies (200 µl per well of a 96 well plate):

Mouse IgG control (Santa Cruz, sc-2025): concentration matched to experimental antibodies

DX2 anti-Fas monoclonal antibody (R&D Systems, mab142), 1/1000 dilution

Anti-HLA Class I antibody (Abcam, W6/32), 1/2000 dilution

Secondary antibodies (100 µl volume per well of a 96 well plate):

Alexa-Fluor-647 goat anti-mouse IgG antibody (Invitrogen, A21237), 1/500 dilution

2.13. Casapase-Glo 3/7 assay

HFFF-hTERTs were seeded in a 96 well plate at $5x10^3$ per well and infected with HCMV Merlin virus at MOI 10, or mock-infected. At 60 hours p.i., culture media were removed and replaced with 50 µl low fluorescence (LF) media containing 10 µl/ml cyclohexamide (Sigma, C6255) and 1µg/ml of one of the following for 12 hours:

-recombinant human Fas ligand (IBA-lifesciences, 2-3901-002)
- recombinant human soluble TRAIL receptor 2 (PeproTech, 310-19)
- Fas mAb (Beckman-Coulter, IM2387)
- mouse IgM isotype control (Sigma, M5909)
negative control wells containing LF media and cyclohexamide only were also included

LF media: Hank's balanced salt solution (no phenol red, 14025)/ 2% FCS/ 2% sodium bicarbonate/ 2% penicillin-streptomycin/ essential amino acids (1/50 dilution, 11130, Invitrogen), non-essential amino acids (1/100 dilution, 11140, Invitrogen).

At 72 hours p.i., Casapase-Glo 3/7 assay (Promega, G8091) reagents were prepared according to the manufacturer's instructions: buffer and lyophilised substrate were equilibrated at room temperature, and buffer added to the substrate. 50 µl of the resulting solution were transferred to each well of the 96 well plate, and samples incubated for 1 hour at 37^{o}C. All samples were then transferred to a white-walled 96 well plate (Thermo scientific, 236105). Blank reactions containing LF media and Casapase-Glo 3/7 assay reagents but no cells were also prepared. Luminescence was measured by the FLUOstar Omega microplate reader in relative light units (RLU).

2.14. Purification of V5-tagged protein

V5-tagged gpUL4 was purified from the culture media of recombinant Merlin-, and RAd-infected cells. For recombinant Merlin-expressed gpUL4, HFFFs were grown to confluency in 3x T150 flasks, infected with RCMV1567 at MOI 1, and maintained in 10 ml complete media per flask. Supernatants were collected 7 days p.i.. For RAd-infected cells, HFFF-hCARs were grown to confluency in 3x T150 flasks, infected with RAd-UL4 at MOI 25 and maintained in 10 ml complete media. Supernatants were collected 3 days p.i. and replaced with fresh media. 4 days later, a second collection was made and cells were re-infected with RAd-UL4 at MOI 50. Supernatants were removed 3 days later, and pooled with the previous two collections.

Soluble gpUL4 was then purified using a purification kit (MBL International Corporation, 3315A) as described by the manufacturer. Briefly, the wash solution was diluted 10 times in distilled water to obtain working concentration. The elution peptide was reconstituted with 100 µl distilled water. Culture media were centrifuged at 1500 rpm for 5 minutes (Megafuge 1.0 R, Heraeus instruments) to remove cell debris, and transferred to fresh 15 ml tubes. Anti-V5 tag beads were resuspended by agitation, and 20 µl of the resulting solution dispensed into each 15 ml aliquot of supernatant. Samples were incubated at 4^{o}C for 1 hour with end-over-end mixing, and centrifuged at 400 rpm for 5 minutes. Supernatants were discarded, pellet of beads resuspended in remaining supernatant (100-400 µl), and transferred to a spin column. Caps and bottom plug were removed from column, which was placed in a collection tube and centrifuged for 10 seconds (Biofuge Fresco, Heraeus). Flow-through was discarded, 200 µl wash solution were added, and the spin column centrifuged for a further 10 seconds. This wash step was repeated two additional times, and the spin column transferred to a fresh collection tube. V5 tag beads were resuspended in 20 µl elution buffer, incubated at room temperature for 30 minutes and tubes centrifuged for 10 seconds. This elution step was repeated once more, except with a 5 minute incubation time. Elutes were pooled together and soluble gpUL4 concentration measured by the Bradford protein assay using the NanoDrop ND-1000 spectrophotometer

3-Cloning and expression of HCMV genes by Using an Adenovirus Vector

3.1. Adenoviruses vectors and HCMV research

HCMV encodes an impressive array of immune-evasion functions, some of which have not been mapped to specific genes (see section 1.7). Discovering novel immune-evasion HCMV functions and mapping genes responsible for known functions can aid our understanding of virus life cycle and pathology, and also inform us of the way the immune system operates. However, the sheer complexity of the virus hinders identification and analysis of gene function. The generation of HCMV mutants harbouring single gene deletions and the production of high-titer HCMV stocks is time and labour intensive, and limited to non-essential genes. In preference to generating an extremely large number of single-gene-mutant viruses, HCMV genes were expressed in isolation as a means to systematically map gene function.

Replication-deficient Recombinant Adenovirus vectors (RAds) are extensively utilised as gene transfer vehicles, and their use in large scale studies of gene function is also increasing. RAds provide a number of advantages as a means of gene delivery vectors, such as *E.coli* -based gene cloning and manipulation technologies, the insertion of large inserts, a broad range of permissive cell types, efficient gene delivery (up to 100% of target cells), and potential for high level gene expression (Wilkinson and Akrigg, 1992). Ad serotype 5 is the best characterised and most commonly used Ad vector in gene therapy. High titer Ad 5 stocks (>10^9 pfu/ml) are readily produced, and the technology of generating recombinants of this virus is well developed. RAd vectors tend to be deleted in both the essential early 1 (E1A), and the 'non-essential' E3 regions of the Ad genome; thus allowing insertion of transgenes up to 8 kb in size. To complement the E1 deletion, Ad vectors must be propagated on E1-expressing helper cells. The E1A gene has been stably inserted into the human kidney cell line 293 (Graham et al., 1977); these 293 cells are well suited to Ad propagation, as they are readily transfectable, and facilitate growth of the virus to high titers. Extremely high level, constitutive expression of recombinant genes under the control of the HCMV major IE promoter was achieved when this promoter was inserted into the Ad 5 genome (Wilkinson and Akrigg, 1992).

The potential to achieve high-level transgene expression from RAds has previously been exploited in studies into HCMV gene function. UL40 expressed from a RAd efficiently upregulated HLA-E expression at the cell surface, and elicited protection from NK cell lysis (Tomasec et al., 2000). The MHC class I-related molecules gpUL18 and UL142 also inhibited NK cells when expressed from RAds (Prod'homme et al., 2007, Wills et al., 2005a). Moreover, RAd-

expressed US6 functionally mimicked US6 produced during productive HCMV infection, by downregulating MHC class I surface expression (Tomasec et al., 2000).

3.1.1. AdZ: a RAd optimised for high-throughput gene cloning and expression

Although RAd technology has yielded great success in testing predicted HCMV gene functions, such studies have generally been carried out sporadically and not on a large scale (Wilkinson et al., 2008). This is mainly because traditional gene cloning and generation of RAds remained laborious, and unsuitable for high-throughput cloning of large numbers of genes. There was therefore a compelling need for an improved RAd cloning and expression system. The recent generation of the AdZ (Ad with zero cloning steps) vector in our laboratory overcame deterrents such as the need for multi-step cloning protocols (Stanton et al., 2008b), thus enabling high-throughput gene cloning by the single step recombineering (recombination-mediated genetic engineering) method (Lee et al., 2001, Yu et al., 2000). Recombineering utilises the λ Red-encoded gene functions expressed from a stably integrated defective λ prophage in the E.coli strain SW102. The three λ Red-encoded genes exo, bet, and gam are under the control of the strong phage promoter pL, which in turn is stringently controlled by the temperature-sensitive repressor cI857 (Warming et al., 2005). λ Red genes are inactive at 32^{o}C, but shifting the temperature of the bacterial culture up to 42^{o}C results in their activation, and efficient homologous DNA recombination. Recombination occurs between DNA sequences with as little as 30 base pairs of homology, and can be used to produce essentially any type of modification in a target BAC, such as insertion, deletion, and point mutation. Apart from being compatible with the recombineering system, the available AdZ vectors have other advantages; transgenes can be cloned in their native form, or directly fused to epitope (strep, V5, hexahistidine) or fluorescent (GFP, RFP) tags. In human cells, the I-*Sce*I homing endonuclease causes the self-excision of BAC sequences, allowing the linearisation step prior to transfection to be bypassed. The presence of tet-operators downstream of the HCMV major IE promoter efficiently suppresses transgene expression during virus replication in 293 helper cells that express the tet repressor (293TREx). This is important as a proportion of transgenes is incompatible with replication of the vector in the helper cell line. Figure 3.1 is a schematic representation of the AdZ vector, detailing its key features.

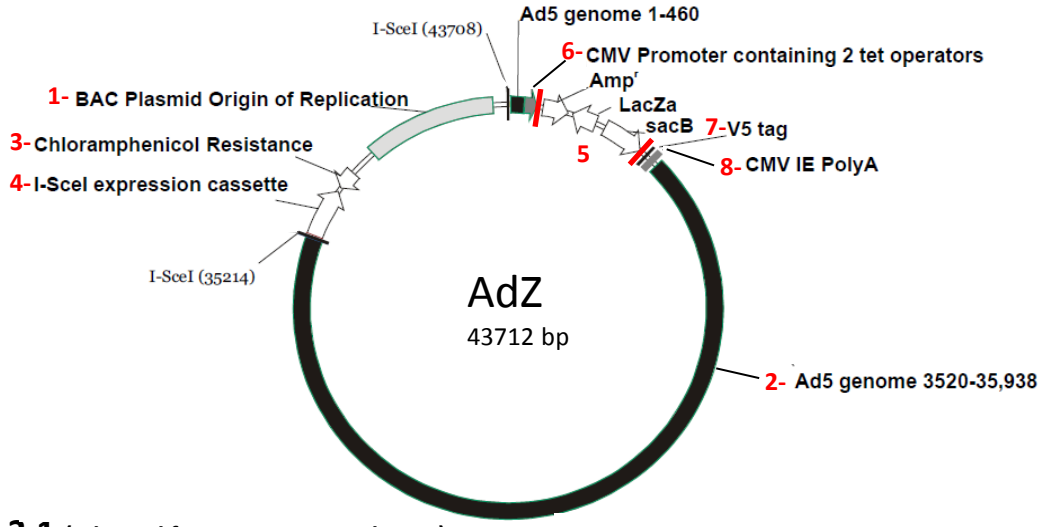


Figure 3.1 (adapted from Stanton et. al 2008):

Scale diagram showing key features of the AdZ vector

1- Origin of replication maintains AdZ as a single copy vector in *E.coli*. 2- E1, E3-deleted Ad5 genome transferred from the AdEasy vector (He et. al . 1998); numbering taken from GenBank sequence AC_000008. 3- Chloramphenicol resistance gene in the plasmid backbone allows selection of AdZ-containing E.coli by chloramphenicol. 4- I-*Sce*I homing endonuclease expression is driven by eukaryote-specific respiratory syncytial virus (RSV) promoter. 18 bp long I-*Sce*I restriction sites are present at the termini of the Ad5 genome; transfection of the AdZ vector into 293TREx cells releases the adenoviral genome in linear form, rendering it infectious. 5- The *ampicilin resistance/LacZ/Sac*B cassette (flanked by red lines) acts as a cloning site; it is positioned in place of the deleted E1 gene, and is replaced by the transgene through recombineering. *Bam*HI restriction sites are also present either side of the cassette, and allow restriction digest analysis. 6- The poweful HCMV MIE promoter (Strain AD169, -299 to +60) constitutively drives trangene expression in permissive cells. This promoter also contains 2 tetracycline (tet) operator sequences, allowing efficient suppression of transgene expression in cells expressing the tet repressor protein. 7- 14 amino acid-long V5 tag is expressed C-terminal to recombinant proteins, allowing expression to be detected with the anti-V5 antibody. 8- polyadenylation signal from the AD169 MIE gene is present in primary transcripts of the transgene.

3.2. Generation of a high-throughput gain of function screening system

With the aim of identifying novel HCMV- encoded immune-modulation functions, and in particular NK-evasion functions, my research group in Cardiff sought to clone every HCMV Merlin gene into the AdZ vector. The library of recombinant adenoviruses would then be used in high-throughput gain-of-function screens. In support of this endeavour, the complete strain Merlin genome had recently been BAC-cloned to provide a reliable source of wildtype HCMV genes (see Section 1.2.2).

3.2.1. Cloning HCMV genes into AdZ

HCMV strain Merlin genes were to be cloned into AdZ by recombineering (see Section 2.7) in frame with a C-terminal V5-epitope tag, allowing their expression to be validated using a single V5 antibody. The C-terminal tag was incorporated into the expression construct primarily as a device to validate expression, even though there was potential for some interference with gene function as a result of the modification. Moreover, specific antibodies are not available for the detection of most HCMV proteins. The V5 tag was chosen based on superior detection; when a group of ten HCMV genes was cloned with hexahistine, Strep, and V5 tags in parallel, expression of all constructs was observed only with the V5 tag (Dr Richard Stanton, unpublished data).

HCMV Merlin genes were individually amplified from the Merlin BAC, and inserted into the AdZ vector. Genes were amplified using primers homologous in their overhanging 5' ends to regions flanking the insertions site, by single run or nested PCR. Primer sequences for genes I cloned are presented in Appendix II. Amplified genes were then introduced into the AdZ vector by recombineering, replacing a cassette of selectable markers. Figure 3.2 illustrates the amplification and cloning of HCMV genes into AdZ. After recombineering, bacteria were recovered and plated on negative selection plates (see Section 2.3.8.1). For each gene, 4 white colonies were inoculated into 5 ml LB/chloramphenicol, and grown overnight. BAC DNA was prepared from the overnight cultures, and *Bam*HI restriction digestion performed. Restriction digestion products were then electrophoresed on an agarose gel. The AdZ vector contains 7 *Bam*-HI restriction sites in its sequence, including 2 either side of the selection cassette, and 1 within. Restriction digestion of AdZ vector DNA with *Bam-HI* generates 18, 11, 7.7, 2.5, 1.7, 0.8 and 0.6 Kbp DNA fragments; the 2.5 and 1.7 Kbp molecules are formed from excision and single-site digestion of the *ampicillin resistance/Lac*Z/*Sac*B cassette. Therefore when a

transgene is inserted in place of the cassette by recombineering, the 2.5 and 1.7 Kbp bands are lost. The presence of a band corresponding to the size of the transgene, and the absence of the 2.5 and 1.7 Kbp bands therefore indicate successful cloning. In these colonies, inserted DNA was sequenced using primers that bind the HCMV promoter and polyA sequences.

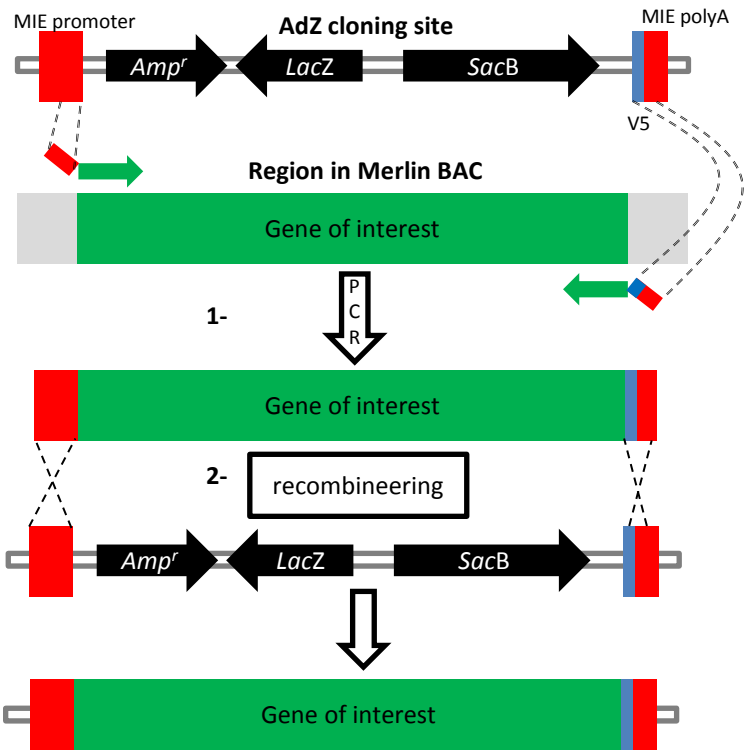


Figure 3.2: Inserting HCMV genes into AdZ by recombineering

1- Genes were amplified from the Merlin BAC using primers homologous in their overhanging 5' arms to regions flanking the selection cassette. 2- PCR products replaced the selection cassette in AdZ through recombineering, and positive recombinants were identified by negative selection. Once the sequence of the insert was verified, BAC DNA containing the transgene was transfected into 293TREx cells and recombinant virus generated

3.2.2. Generation, propagation and titration of RAds

Once the sequence of the insert had been verified, recombinant DNA was maxiprepared and transfected into 293TREx helper cells. Virus was allowed to propagate, and collected once complete destruction of the monolayer was observed. To obtain high-titer stocks, this initial virus collection was used to infect cells in larger 150 cm^2 flasks, and the subsequent virus collection purified on a CsCl gradient, and dialysed to remove toxic Cs. Each virus stock was then titrated by the spot assay (Bewig and Schmidt, 2000). In the spot assay, virus particles are quantified immunocytochemically, using a primary antibody directed against virus proteins, and HRP-conjugated secondary antibody. Addition of 3, 3'-diaminobenzidine (DAB) as substrate then results in production of a brown precipitate in infected cells, easily visualised by light microscopy as brown spots. In an independent assessment, a subset of viruses was also titrated by the picogreen fluorescence method. When a conversion factor was deployed, comparable values were obtained for quantification of virus particles/infectious units by the spot assay and picogreen fluorescence (data not shown). Picogreen is a fluorescent dye that displays up to 1000 times enhanced fluorescence upon binding to double stranded DNA. A sample of the purified virus preparation was lysed in a solution containing the picogreen dye, and the level of fluorescence measured and compared to that of a reference stock of known titre. Titration by picogreen staining effectively generates a total particle count, because picogreen indiscriminately binds DNA of infectious and non-infectious particles. Using the spot assay however, only infectious particles are accounted for; this is an advantage because virus preparations may contain a variable proportion of inactive particles still containing DNA. It was decided therefore that titers obtained by the spot assay would be used as benchmark.

3.3. Library of Ad recombinants containing HCMV genes

Generation of the RAd-library of HCMV genes was a strategic project undertaken by the Molecular Virology Laboratory involving many researchers. The project was initiated with the cloning of all the UL/b' genes into the AdEasy 1 vector (He et al., 1998) using a C-terminal Strep tag. Both the vector and the epitope tag were not efficient enough to undertake cloning of all HCMV genes. The AdZ vector was generated specifically to enable the rapid cloning of HCMV genes, and the V5 tag deployed in preference. My contribution towards generation of the

library of RAd-encoded HCMV genes included (i) cloning 34 genes into the AdZ vector, (ii) generating 85 recombinant viruses by transfection, and (iii) growing, purifying and titrating 125 viruses. Table 3.1 details the individual contribution made by members of our laboratory to this collaborative project.

To date, 169 genes encoded by strain Merlin have been cloned into Ad vectors, and titrated recombinant viruses produced. These include 165 protein-coding genes originally annotated in strain Merlin (Dolan et al., 2004), as well as the β2.7 gene (possibly non-protein-coding), and newly discovered protein-coding genes RL8A, RL9A and US33A (Gatherer et al., 2011). The protein products of all genes are expressed with C-terminal V5 tags, excepting genes UL111A and UL141 were Strep-tagged while the UL14, UL122 and UL123 constructs were not tagged (specific monoclonal antibodies were available). Barring the UL141 gene that was previously inserted into the AdEasy1 vector, all genes were inserted into AdZ. Newly identified genes UL74A and UL150A were not cloned; these are predicted to be alternatively spliced from multiple upstream exons, and thus present a special challenge (see Section 5). The majority of genes were cloned into Ad vectors and expression of their gene products detected without issue. Problems arising with specific genes are described below.

3.3.1. Cloning of genes resistant to recombineering

A group of 15 genes was identified as being resistant to insertion into AdZ: UL47, UL48, UL54, UL55, UL57, UL70, UL86, UL87, UL89, UL104, UL105, UL131A, UL133, UL136, and UL150. While the *ampicillin resistance/Lac*Z/*Sac*B cassette was removed from AdZ, it was not replaced with the gene of interest. To overcome this shortcoming, cloning of these genes was undertaken by Dr J.Davies using the *gal*K selection method (Warming et al., 2005). The *E.coli* galactose operon is made up of 4 genes, including *galK* that are indispensible for galactose metabolism. SW102 bacteria harbour a precise 1231 bp deletion of *galK* and therefore cannot grow on media containing galactose as the sole carbon source. *galK* naturally phosphorylates galactose to galactose 1-phosphate, but also efficiently phosphorylates 2-deoxygalactose (DOG), a galactose analogue, into a toxic product. The galK gene can therefore be selected for on minimal media containing galactose, and selected against on minimal media containing DOG, and glycerol as the carbon source.

galK was amplified using primers homologous in their 5' ends to a site in the middle of the target gene, and inserted into the Merlin BAC by recombineering. Recombinants were identified by positive selection, BAC DNA was isolated, and the HCMV gene containing within it the galK gene amplified by PCR. A second recombineering round was carried out to insert the purified PCR product into AdZ, and selection again performed for *galK*. In a third round of

recombineering, *galk* was removed from AdZ to leave behind the intact transgene, and recombinants selected against on media containing DOG. The fact that recombineering by galK selection was successful for the above genes argues their sequences were not completely prohibitive of recombineering, but that the recombination event was extremely inefficient.

Some of the aforementioned troublesome genes are amongst the largest in the HCMV genome, and recombineering is known to be less efficient with larger genes. However this is unlikely to completely account for difficulties in cloning these genes, as other large genes were cloned with relative ease. One possibility for this peculiar finding is the formation of complex secondary structures in these genes that are incompatible or interfere with recombination. The average guanosine-cytosine (GC) content in these genes is higher than that of the Merlin genome as a whole (62% compared to 57%). It is well known that high GC content can bring about the formation of secondary DNA structures such as stable hairpins or loops, and it has recently been suggested that these sequences may also prohibit successful recombineering (Nelms and Labosky, 2011).

3.4. Expression analysis of the RAd-library

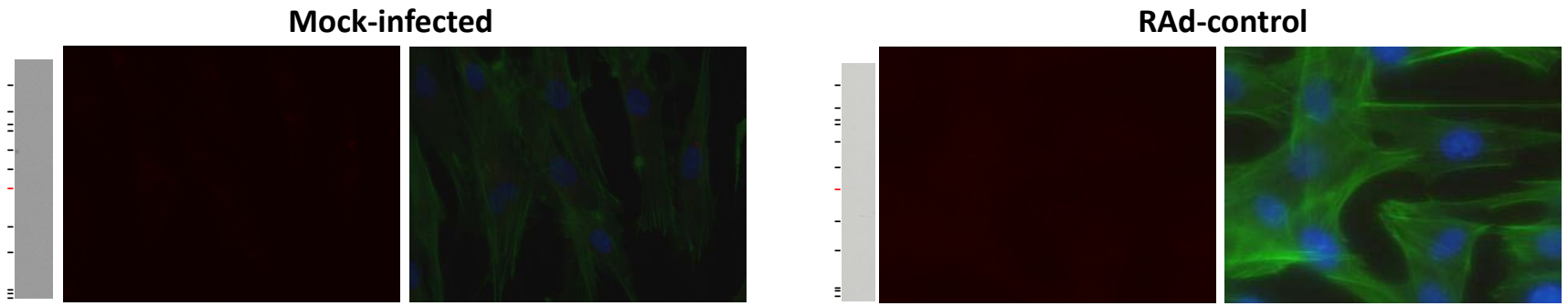
Once titrated stocks were obtained, it was necessary to validate expression of the cloned transgenes in order to provide a secure basis for progressing to functional screening programmes. Although not the primary objective, this analysis is invaluable as it has the capacity to inform on the gene products of all HCMV genes. Expression from RAds in infected HFFF-hCARs was tested by immunofluorescence and western blot using the anti-V5 tag antibody. These assays generated a large dataset detailing the intracellular localisation, apparent molecular mass, and evidence of post-translational modifications of expressed proteins. Of the 169 genes cloned into RAds, expression was confirmed for 160 by a combination of western blot (152 genes) and immunofluorescence (155 genes). These expression data are presented in Figure 3.3, and described in tabular form in table 3.2. Table 3.2 also compares the expression data obtained with previous findings reported in the literature. A detailed evaluation of the expression of all HCMV gene products lies beyond the aims of this study. Nevertheless chapters 4 and 6 contain examples of genes whose functions have been tested and compared in the contexts of RAd and HCMV infection, and this subject is discussed in more detail in chapter 7.

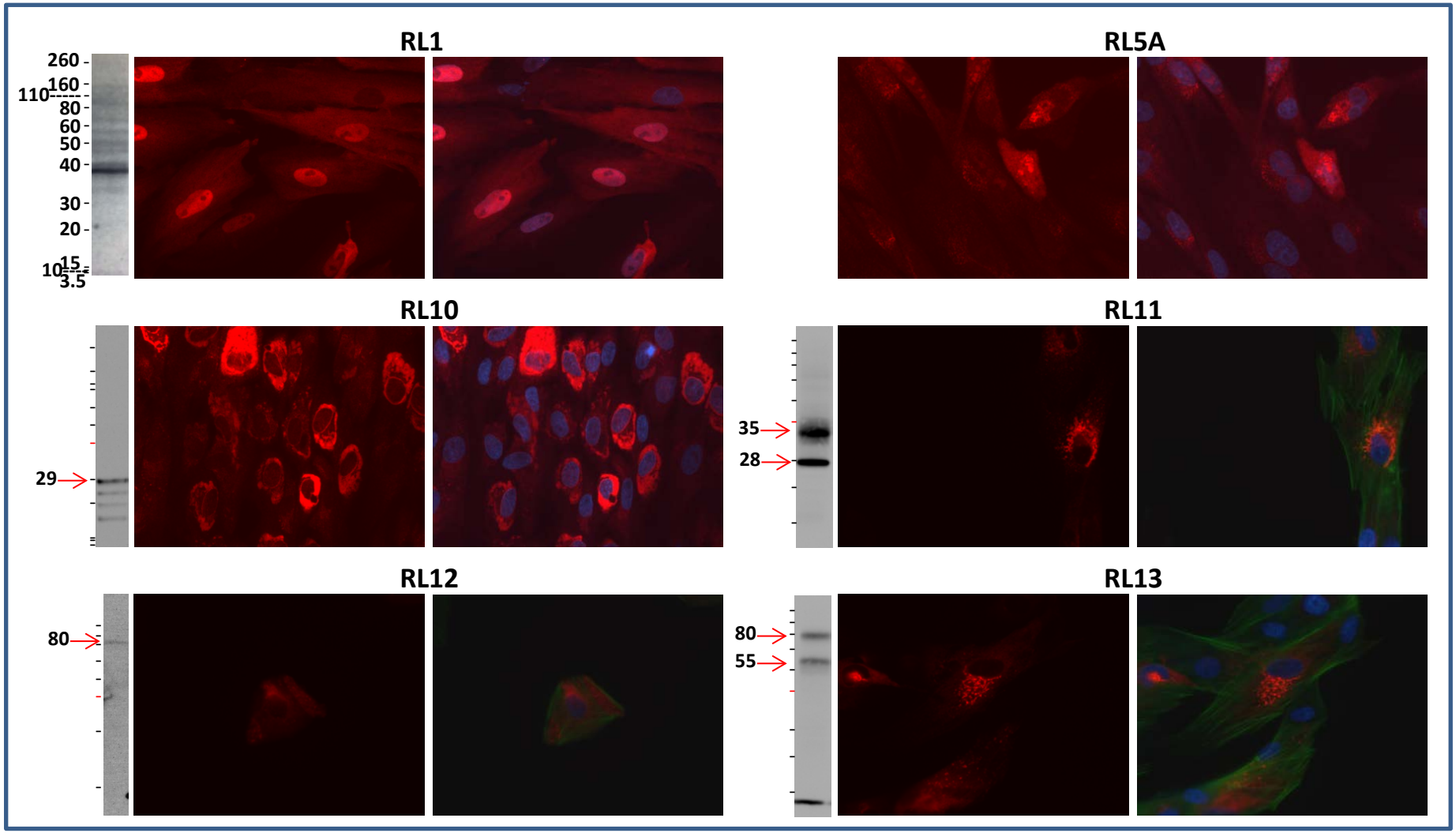
Table 3.1: Contribution of lab members in generating the Ad library of HCMV genes, and immunofluorescence analysis of its expression

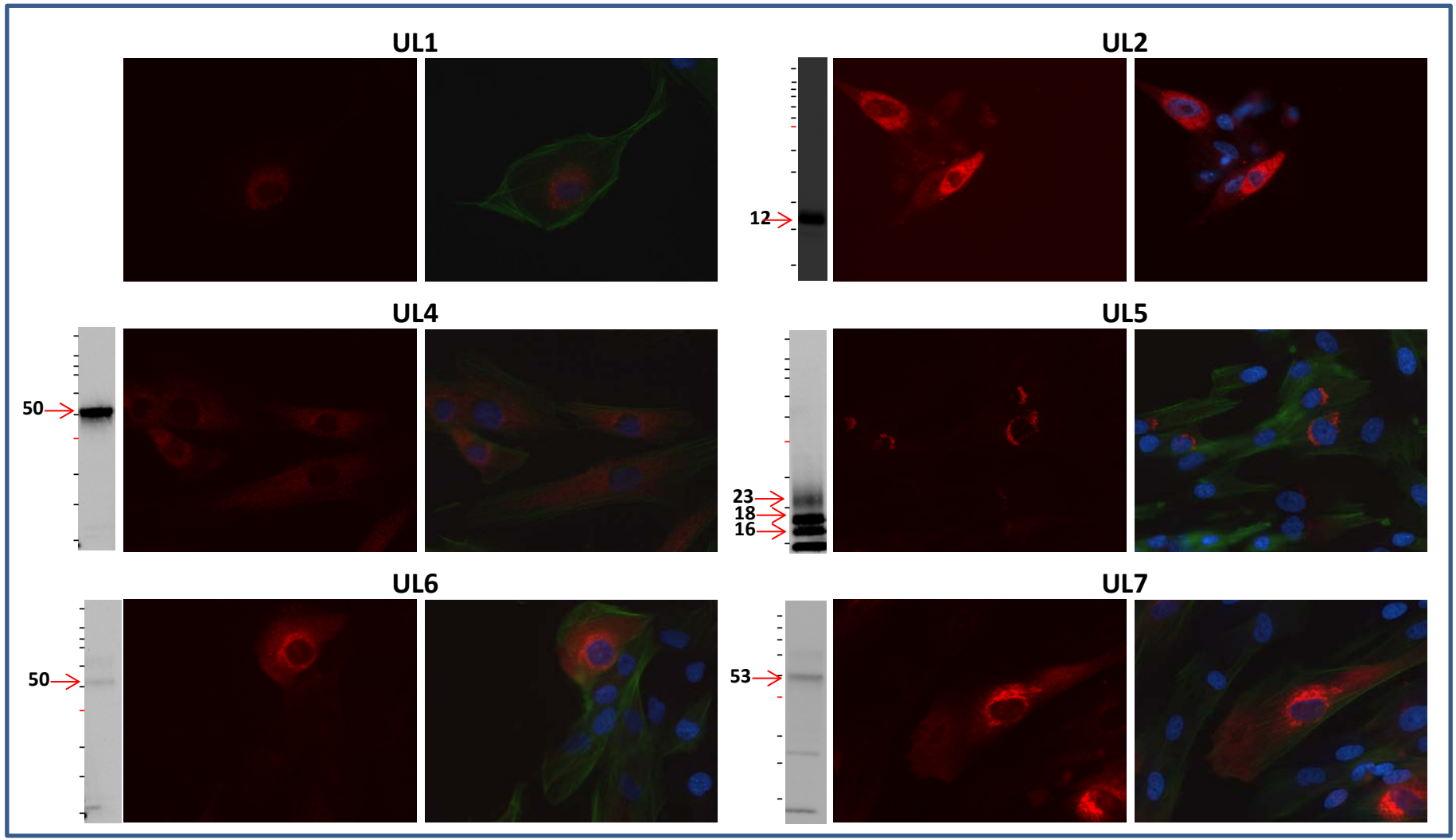
	Cloned into AdZ/**AdEasy** vector	Generated Ad recombinant	Grew, purified, and prepared titered stocks	Performed Immunofluorescence analysis
Dr Richard Stanton	UL111A-UL124 (12 total)	UL111A-UL124 (12 total)	N/A	UL135
Dr Richard Stanton Dr Virginie Prod'homme Dr Song Han	RL1-RL6, RL10-UL6, UL9-UL19, UL22A-UL24, UL30, UL30A, UL35, UL40-UL42, UL44, UL46, UL135, **UL141**, US1-US12, US14, US16 (48 total)	RL1-RL6, RL10-UL6, UL9-UL19, UL22A-UL24, UL30, UL30A, UL35, UL40-UL42, UL44, UL46, UL135, UL141, US1-US12, US14, US16 (48 total)	N/A	N/A
Dr James Davies	UL7, UL20, UL21A, UL25-UL28/29, UL31-UL34, UL37, UL38, UL43, UL45, UL47-UL49, UL51-UL57, UL70-UL72, UL74-UL76, UL78-UL80.5, UL86-UL92, UL94, UL96-UL105, UL131A, UL133, UL136, UL150, IRS1, US13, US15, US17-US22, US26-US32, US34-TRS1 (75 total)	UL7, UL20, UL21A, UL26-UL28/29, UL31, UL33, UL34, UL37, UL38, UL43, UL54, UL55, UL86, UL87, UL89, UL90, UL104, UL105, US13, US17-US21 (24 total)	RL1-RL6, RL10-UL7, UL9, UL13, UL15A- UL31, UL33-UL35, UL37-UL40, UL42-UL44, UL46, UL48A (44 total)	N/A
Sepehr Seirafian	RL8A, RL9A, UL8, UL36, UL50, UL69, UL73, UL77, UL82-UL85, UL93, UL95, UL128, UL130, UL132, UL138-UL140, UL142-UL148D, US23, US24, and US33A (34 total)	RL8A, RL9A, UL8, UL14, UL25, UL32, UL36, UL45, UL47-UL53, UL56-UL85, UL88, UL91-UL103, UL128-IRS1, US15, US22 -TRS1 (85 total)	RL8A, RL9A, UL8, UL10, UL11, UL14, UL32, UL36, UL41A, UL45, UL47, UL48, UL49-TRS1 (125 total)	RL1, RL12, UL1, UL10, UL14, UL18, UL26, UL47, UL48A, UL53, UL111A, UL122, UL123, US34A (14 total)
Daniel Sugrue	N/A	N/A	N/A	UL41A-UL46, UL48, UL49-UL52, UL54-UL105, and US2-TRS1 (77 total)
Sepehr Seirafian (infections and staining) and Daniel Sugrue (imaging)	N/A	N/A	N/A	RL5A, RL10, RL11, RL13, UL2-UL9, UL11, UL13, UL15A-UL17, UL19-UL25, UL27-UL40, UL112, UL121, UL124-UL133, UL136-UL140, UL144-UL148 (62 total)
Melanie Armstrong, PhD thesis, figure 4.3v	N/A	N/A	N/A	UL141

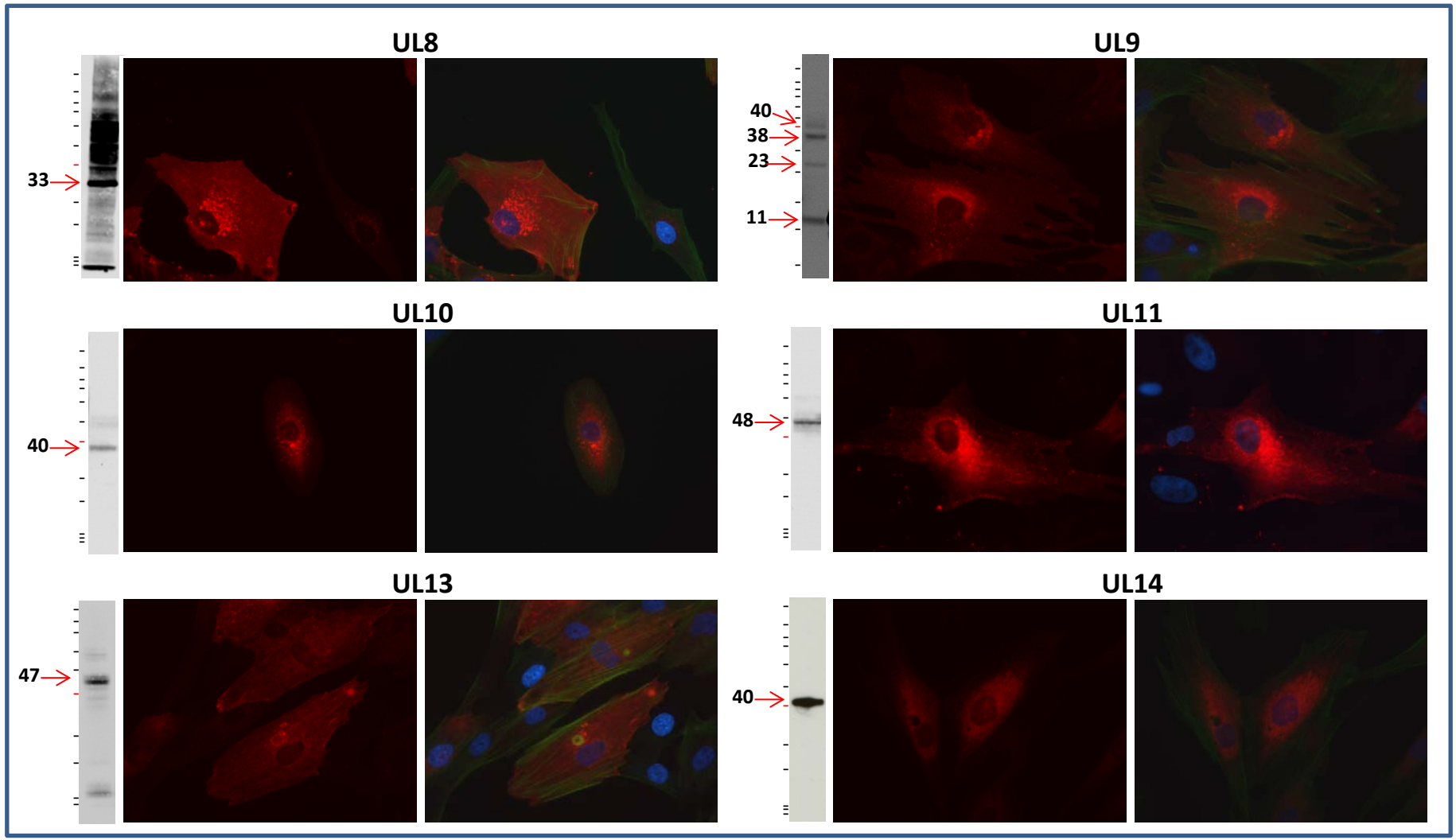
Figure 3.3: Expression analyses of HCMV genes inserted into Ad vectors by Immunofluorescence and Western blot

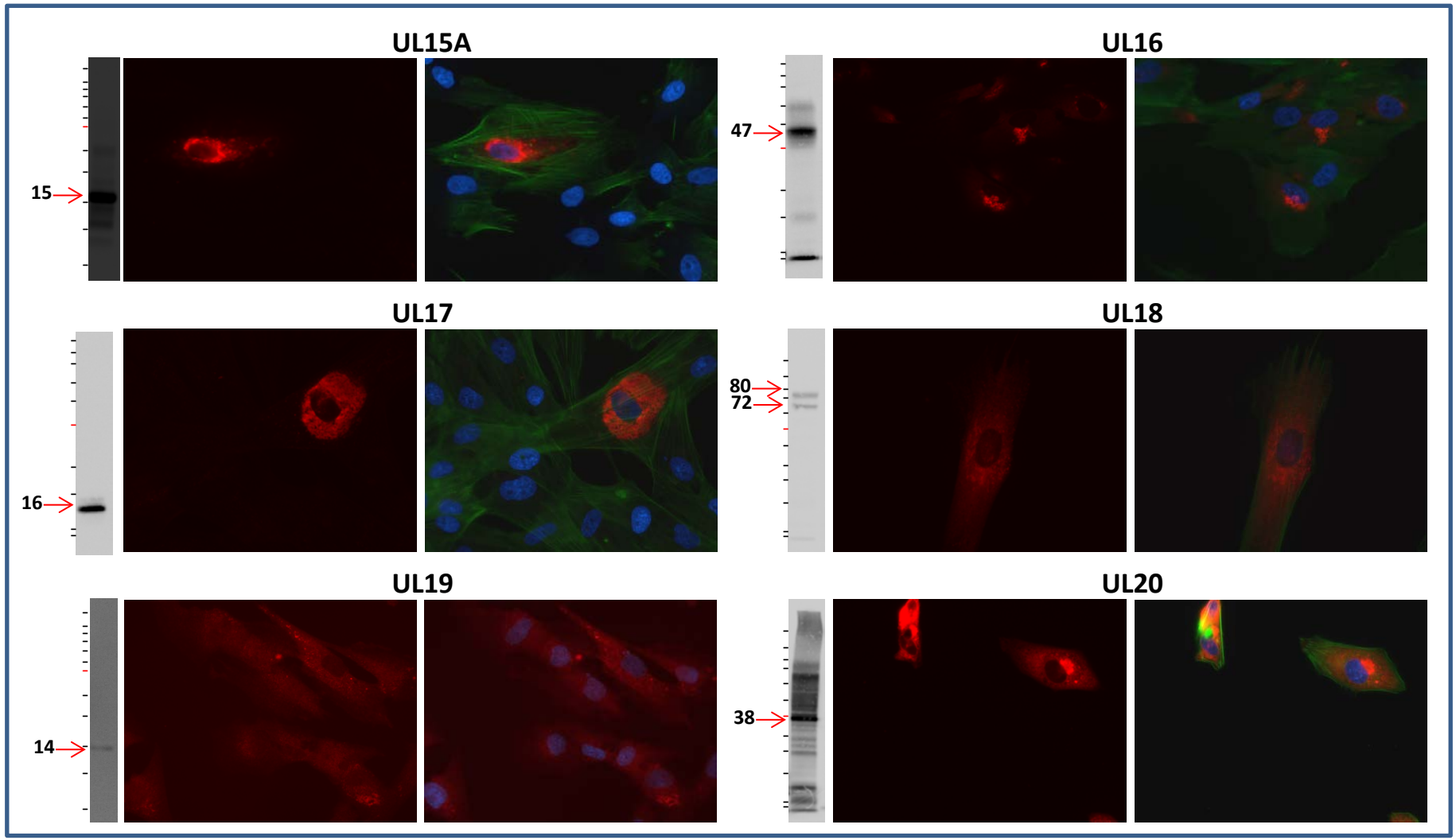
For immunofluorescence, HFFF-hCARs were infected at MOIs 1, 5, 10 and 20 with RAds encoding HCMV genes. Exceptions were RAds encoding RL12, UL1, UL10, UL14, UL18, UL52; these were used at MOI 40. Infection at a range of dilutions was necessary as the level of protein expression varies between different Ad recombinants. 3 days p.i., cells were fixed and permeabilised, and the appropriate primary antibody applied against recombinant HCMV proteins. Monoclonal antibodies were used to detect UL14, UL122 and UL123 gene expression. Expression from genes UL111A, UL135 and UL141 was detected using a Strep antibody, and a V5 antibody was used for all other RAds. An Alexa fluor594 secondary antibody was then added together with DAPI and phalloidin-AF488 which stained nuclei and actin filaments respectively. Immunofluorescence was visualised microscopically using a x40 oil objective. RAd expression (red) is shown alone in the left hand picture, and where available overlaid with DAPI and/or phalloidin-488 staining in the right hand picture. Images shown are those adjudged to be most representative of the whole population, and where expression was observed at different MOIs, the lowest was chosen in order to obtain a closer representation of physiological-level expression**.** Some of the images were provided by other members of our laboratory and are included in order to provide a more complete reference of HCMV gene expression (see Table 3.1). Images presented in appendix III are entirely my own work, but are not included in this figure as they lack co-staining with DAPI and/or phalloidin. For Western blot, HFFF-hCARs were infected at MOI 10. 3 days p.i., infected cells were lysed, and electrophoresed under denaturing and reducing conditions alongside pre-stained protein markers on pre-cast polyacrylamide gels. Following Western transfer, membranes were probed with V5 primary and HRP-conjugated anti-mouse secondary antibodies, except Strep-tagged UL111A and UL135 proteins were detected using Strep antibody, and monoclonal antibodies were used to detect UL14, UL122, UL123 and UL141 gene expression. The sizes of all protein markers are indicated for the RL1 blot as a reference; for all other samples the 40 kDa band is highlighted red and the size(s) of the predominant species indicated by red arrows. No sizes are indicated if single band(s) could not be discerned. The US18 blot was provided by Dr Ceri Fielding. For both techniques, mock-infected cells, and RAd-control (generated from AdZ not containing a transgene)- infected cells were used as negative controls; representative samples are shown below.

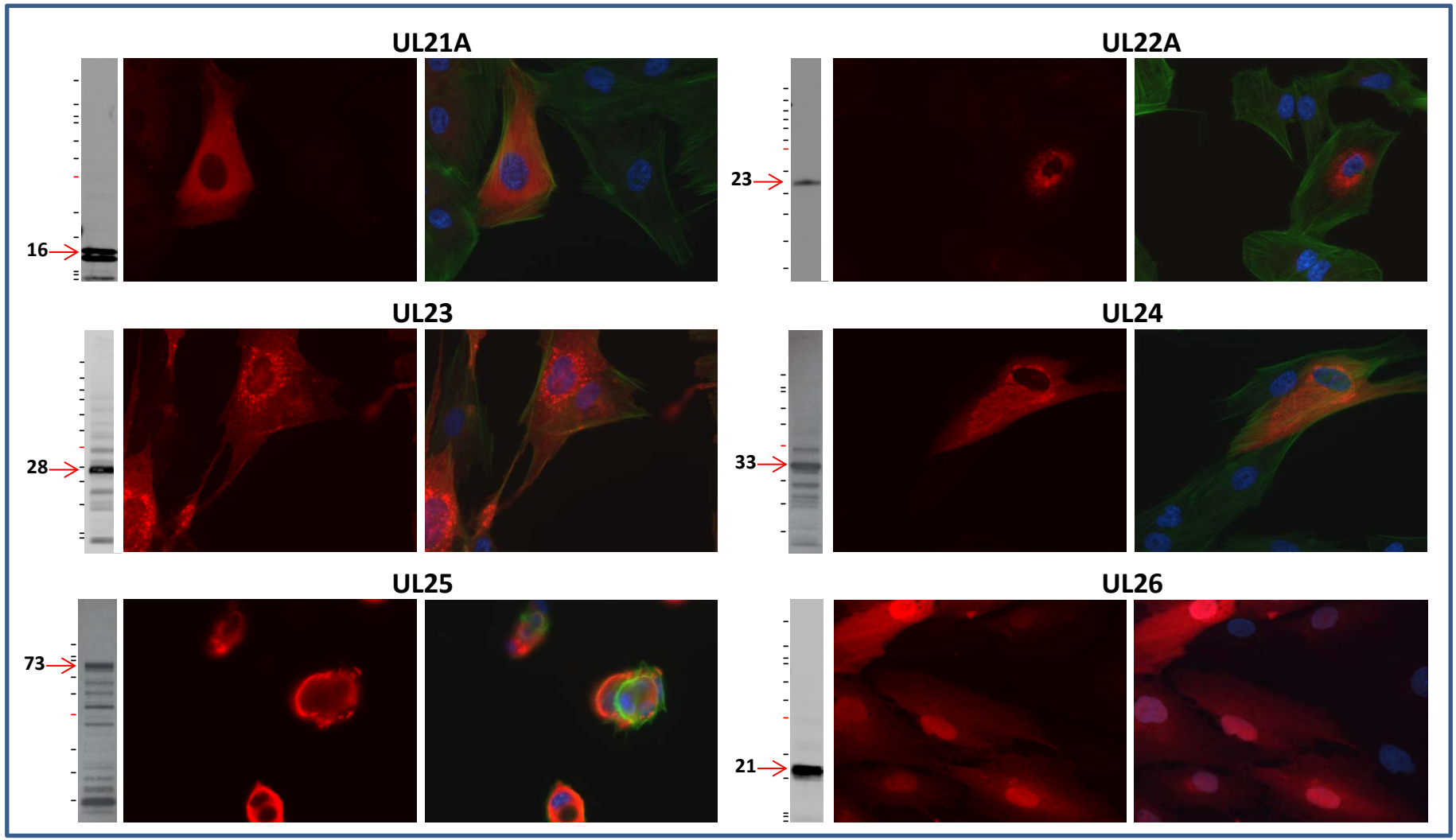


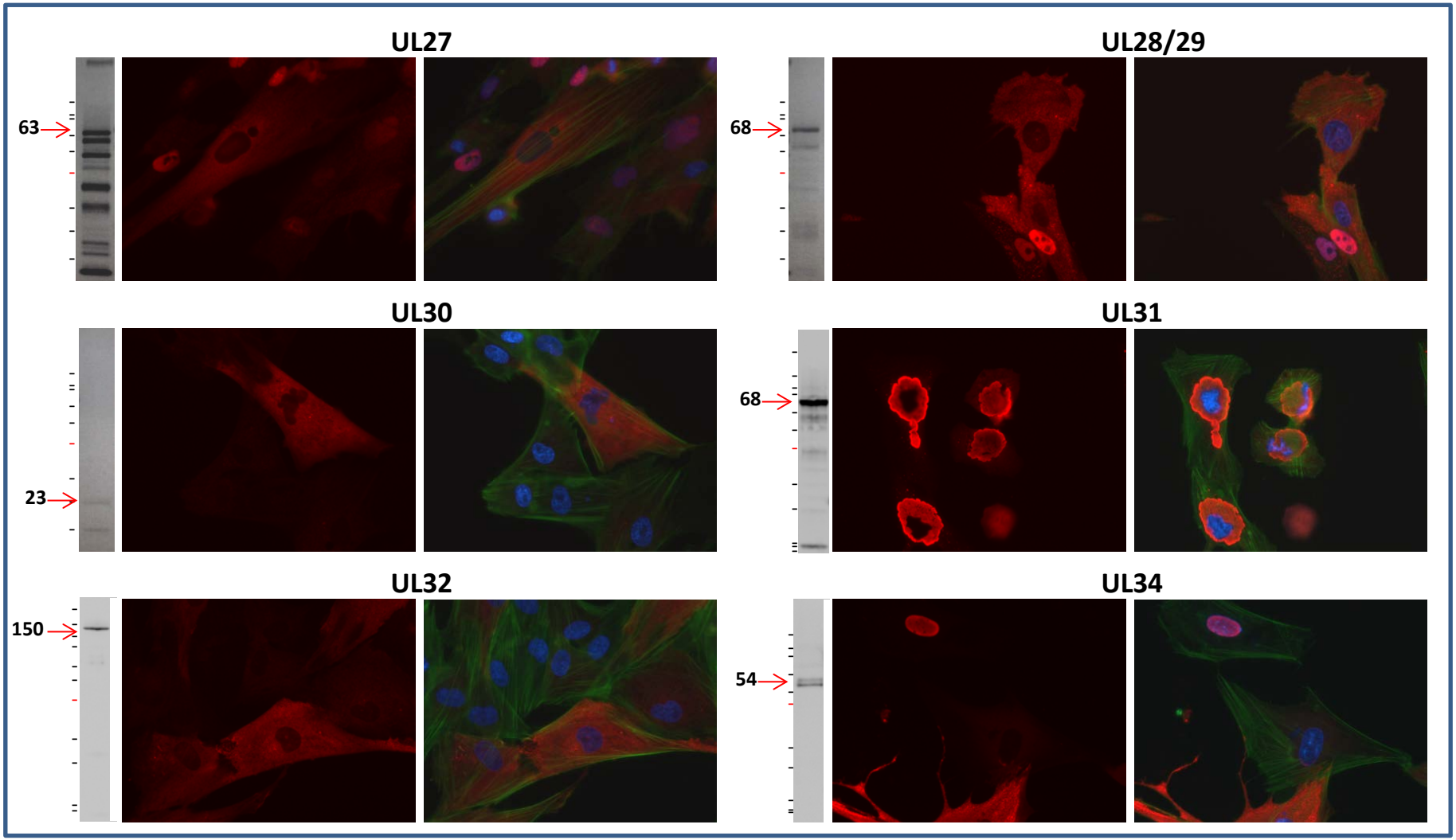

RL1
260
160
110
80
60
50
40
30
20
15
10
3.5
RL5A
RL10
29
RL11
35
28
RL12
80
RL13
80
55

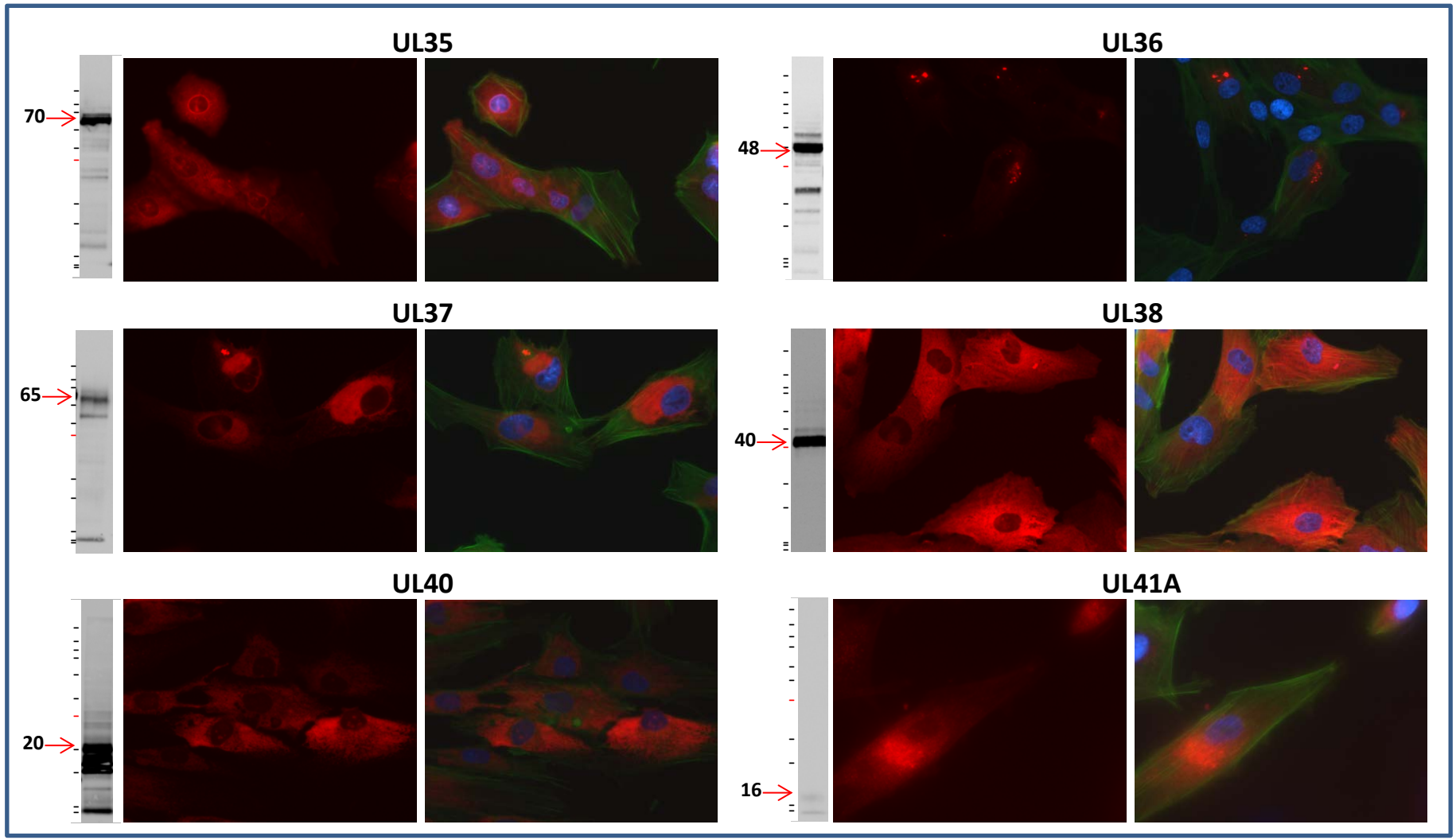

UL1
UL2
12
UL4
50
UL5
23
18
16
UL6
50
UL7
53

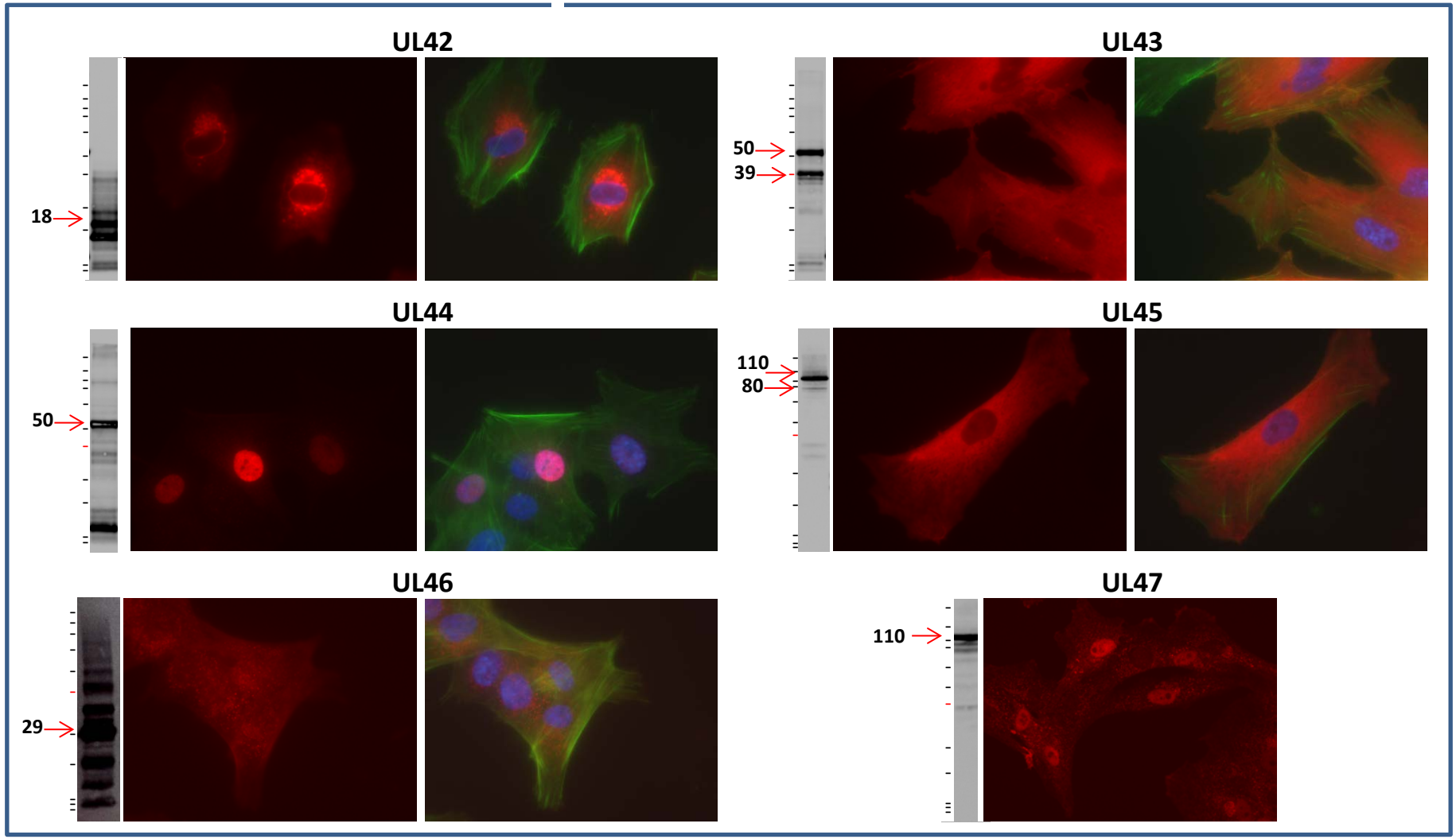

UL8
33
UL9
40
38
23
11
UL10
40
UL11
48
UL13
47
UL14
40

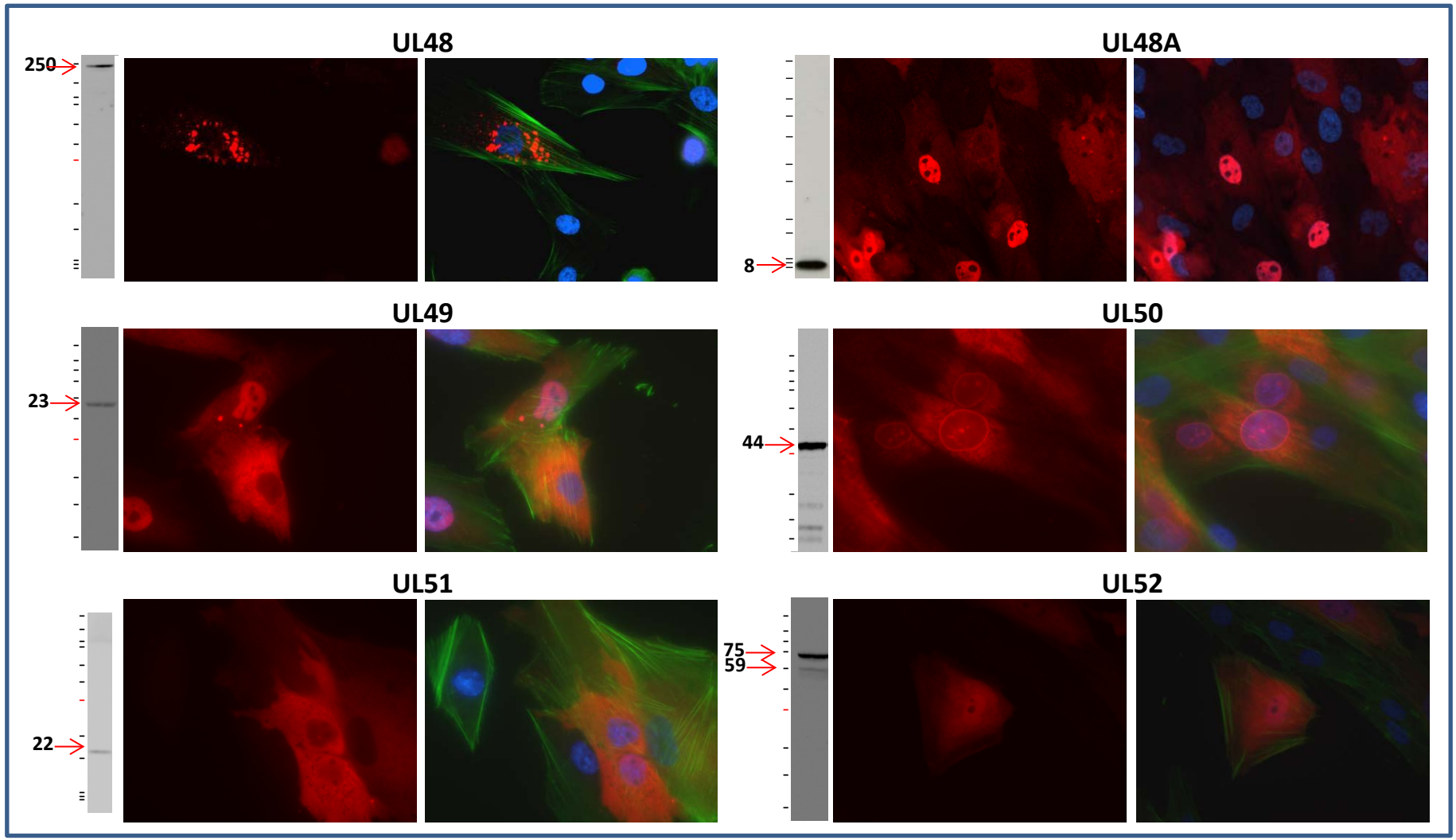

UL15A
15
UL16
47
UL17
16
UL18
80
72
UL19
14
UL20
38

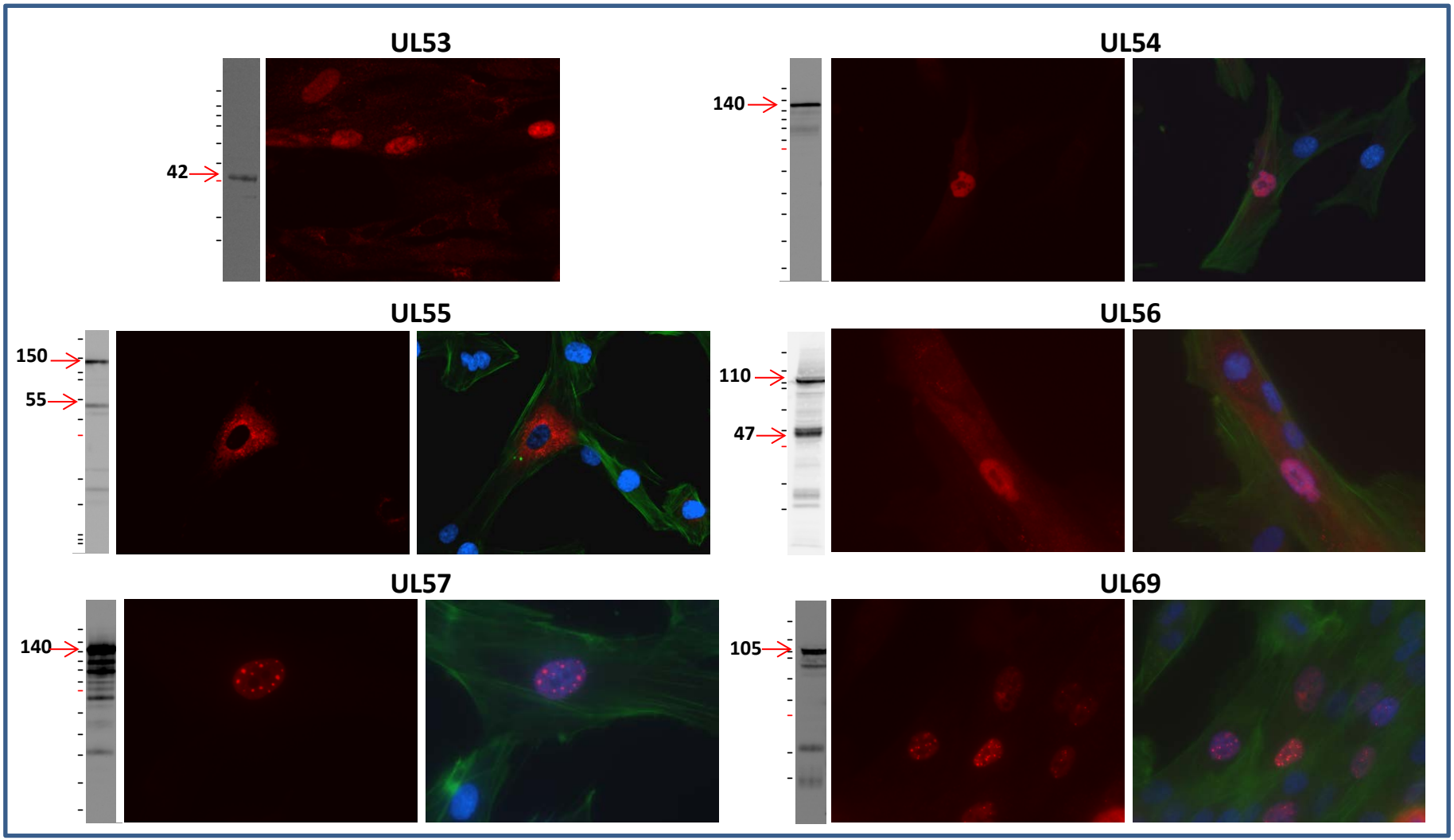

UL21A
16
UL22A
23
UL23
28
UL24
33
UL25
73
UL26
21

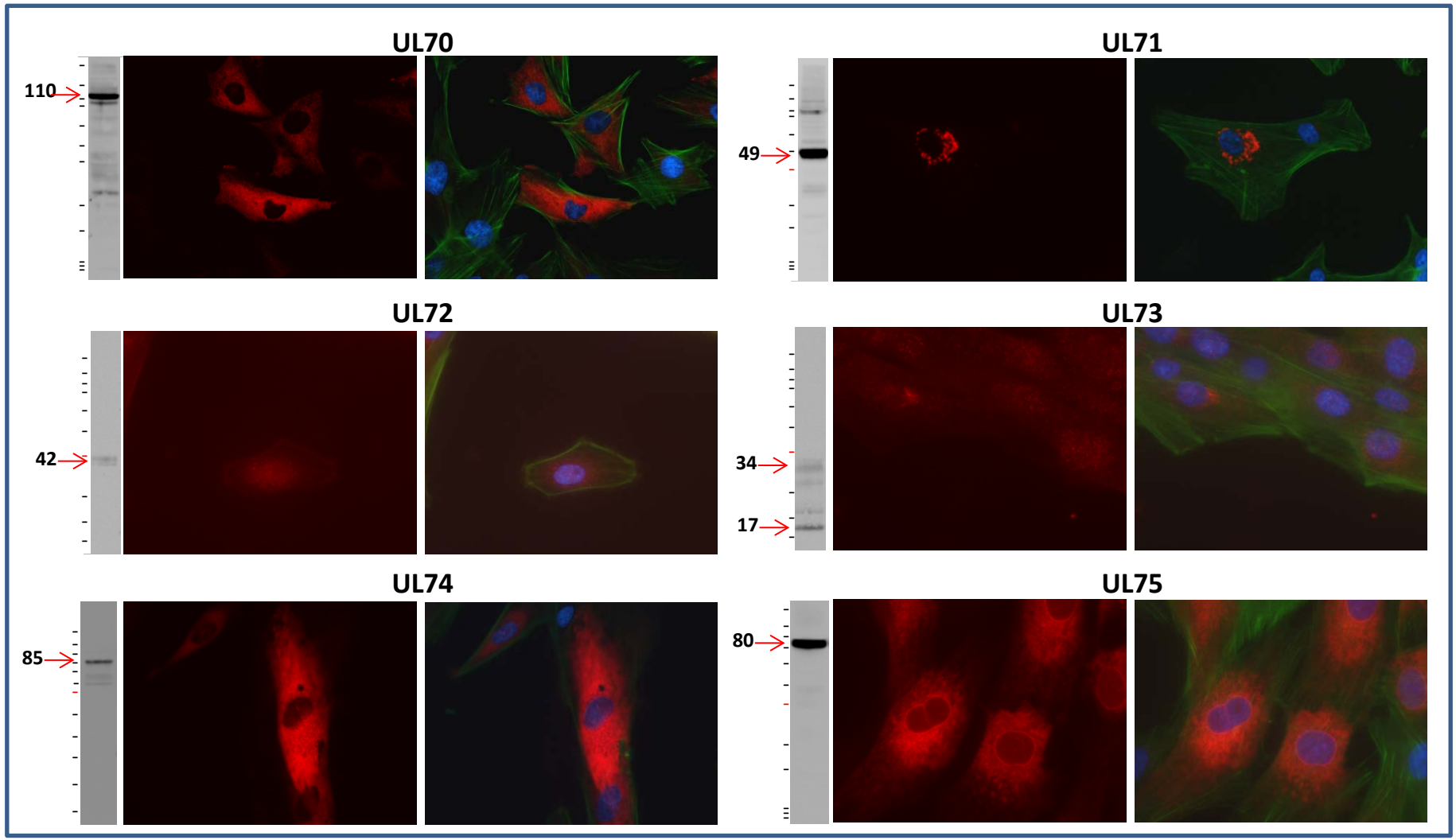

UL27
63
UL28/29
68
UL30
23
UL31
68
UL32
150
UL34
54

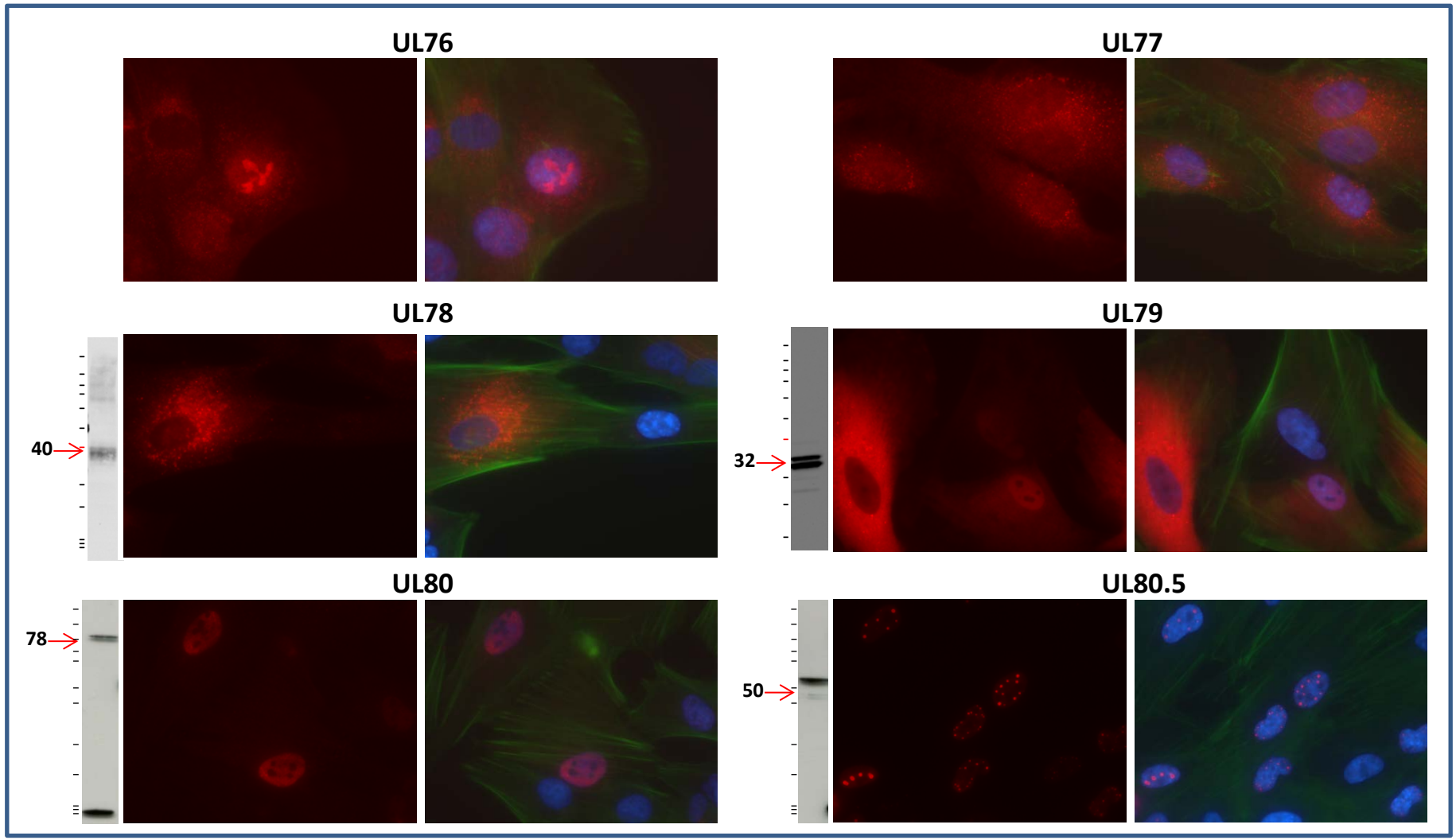

UL35
70
UL36
48
UL37
65
UL38
40
UL40
20
UL41A
16

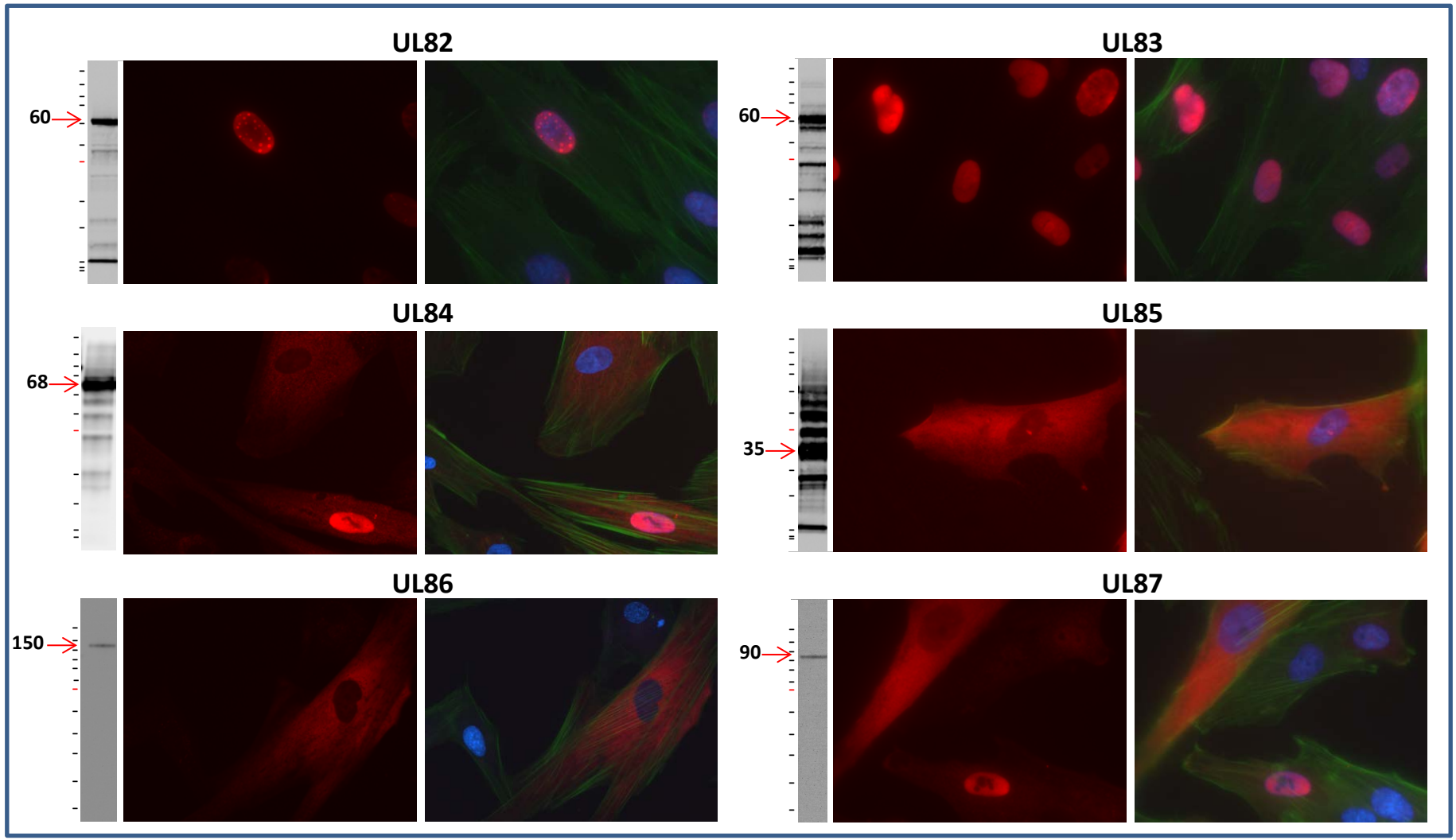

UL42
18
UL43
50
39
UL44
50
UL45
110
80
UL46
29
UL47
110

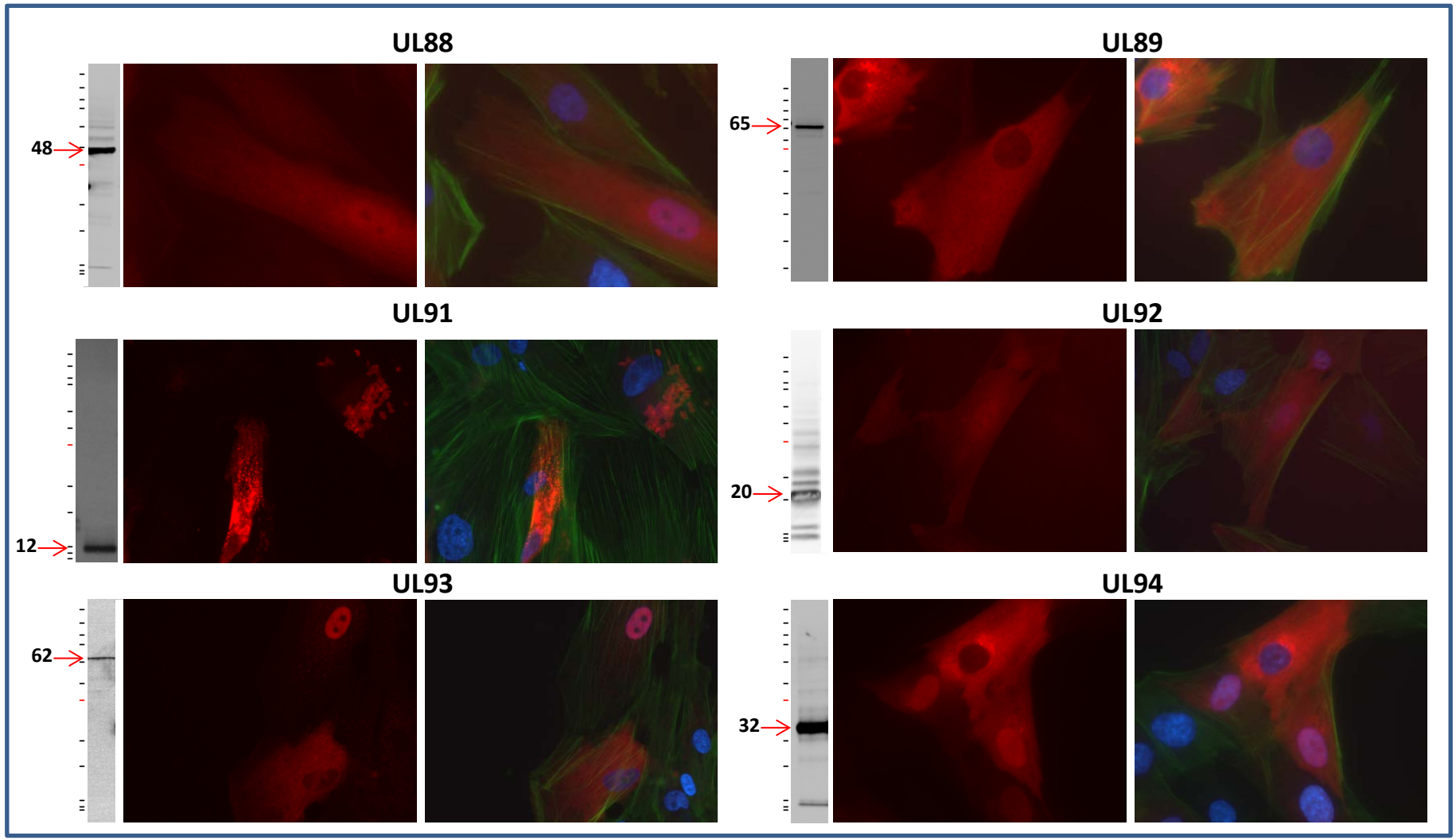

UL48
250
UL48A
8
UL49
23
UL50
44
UL51
22
UL52
75
59

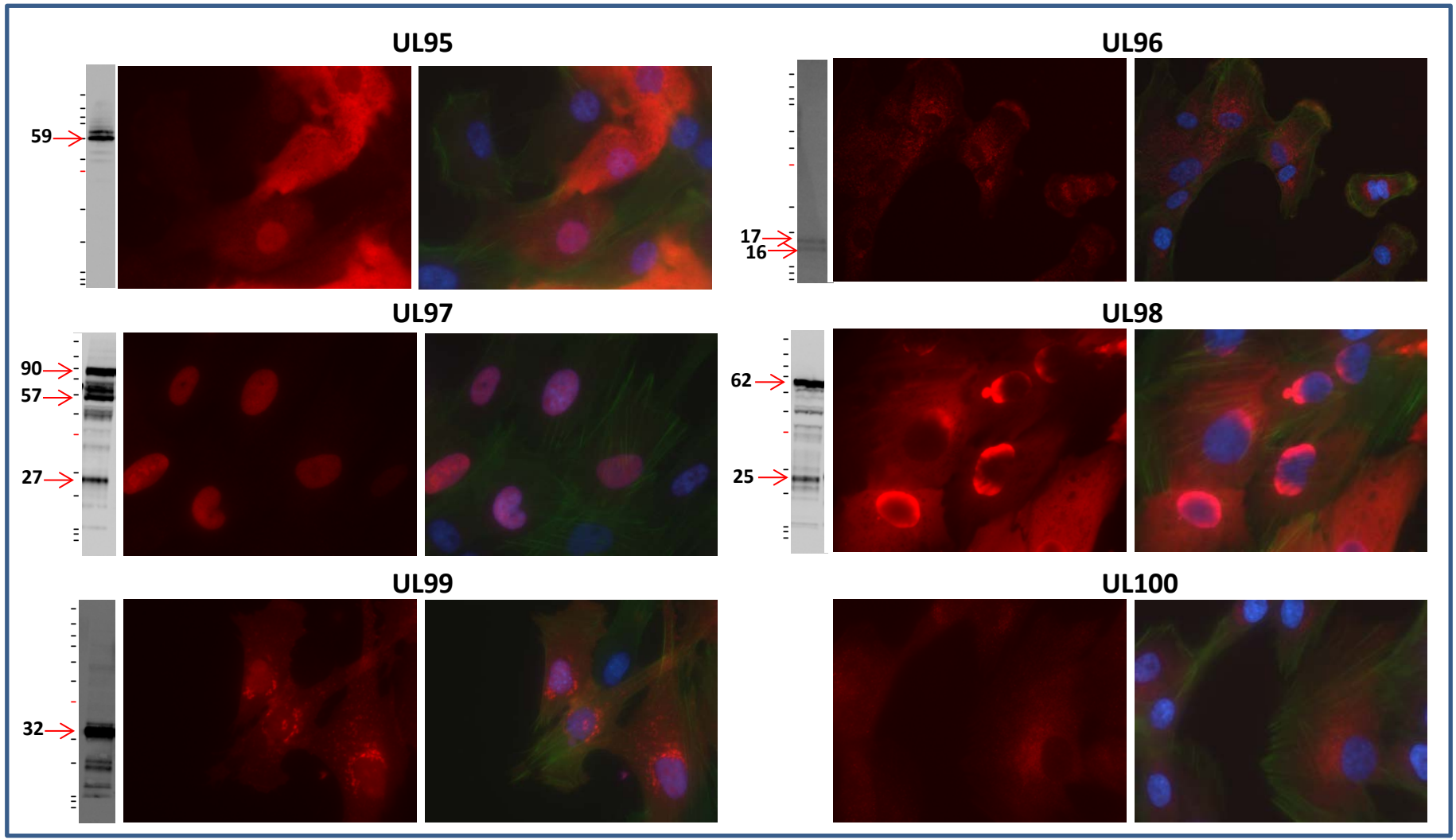

UL53
42
UL54
140
UL55
150
55
UL56
110
47
UL57
140
UL69
105

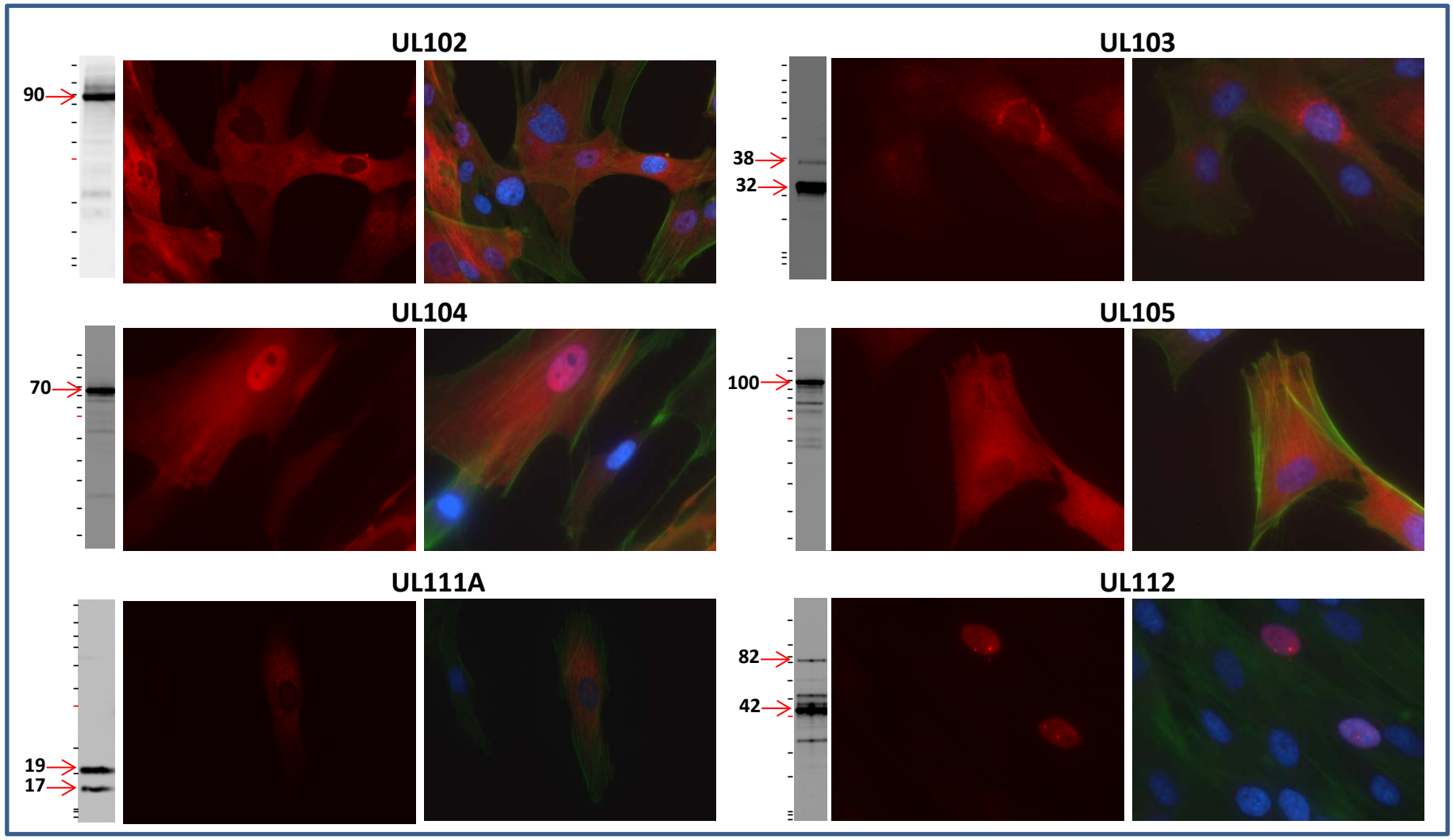

UL70
110
UL71
49
UL72
42
UL73
34
17
UL74
85
UL75
80

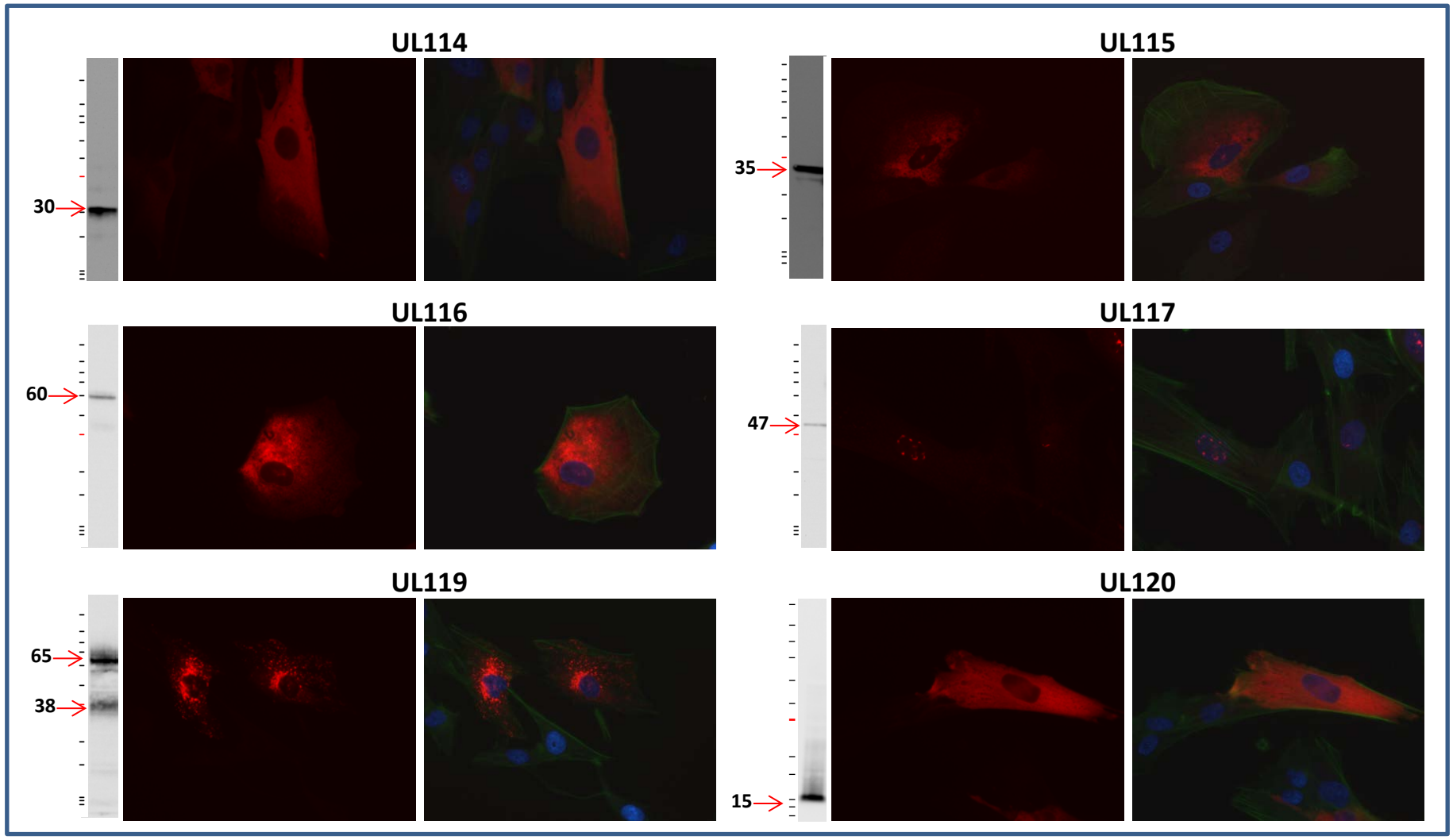

UL76
UL77
UL78
40
UL79
32
UL80
78
UL80.5
50

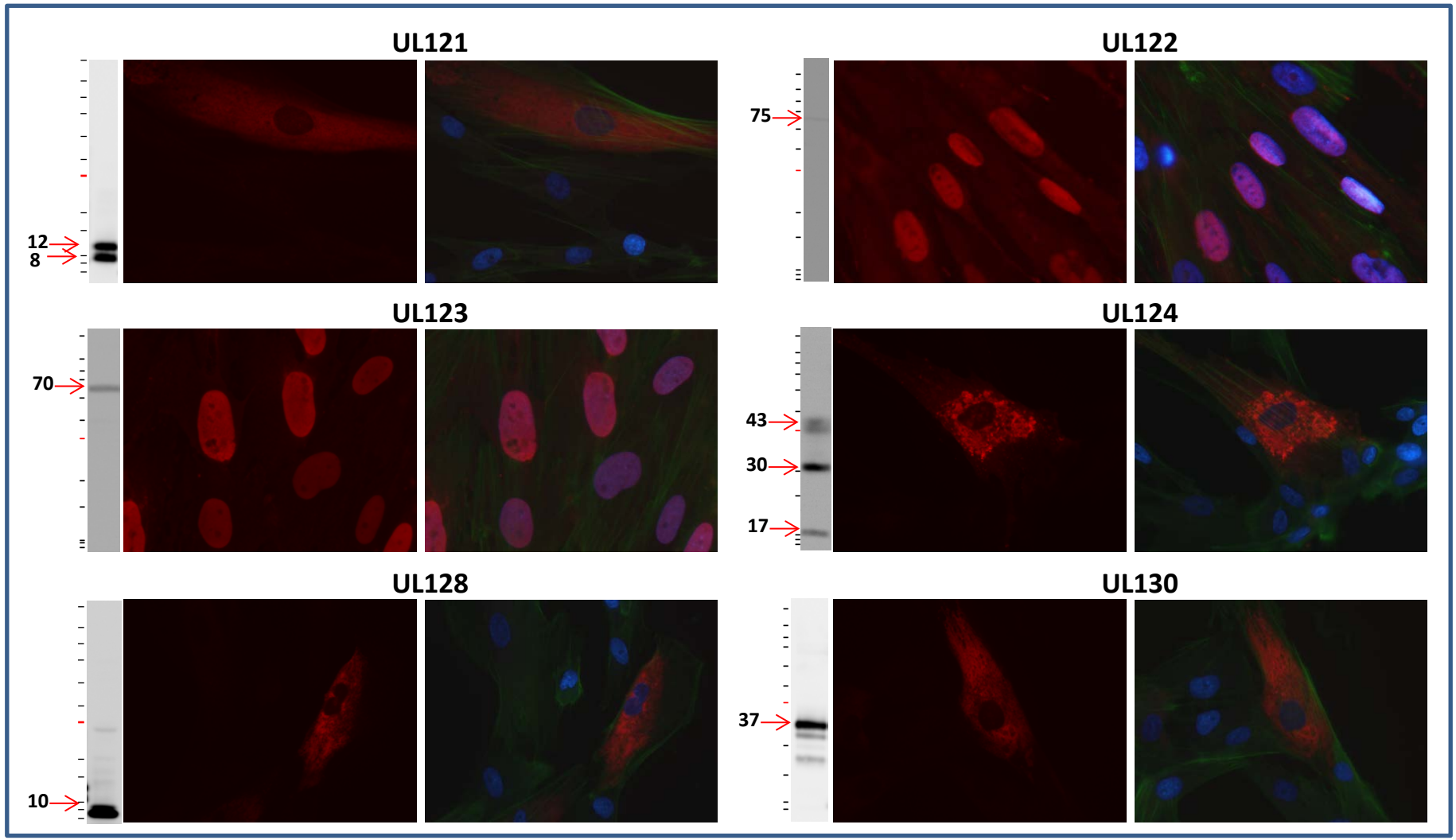

UL82
60
UL83
60
UL84
68
UL85
35
UL86
150
UL87
90

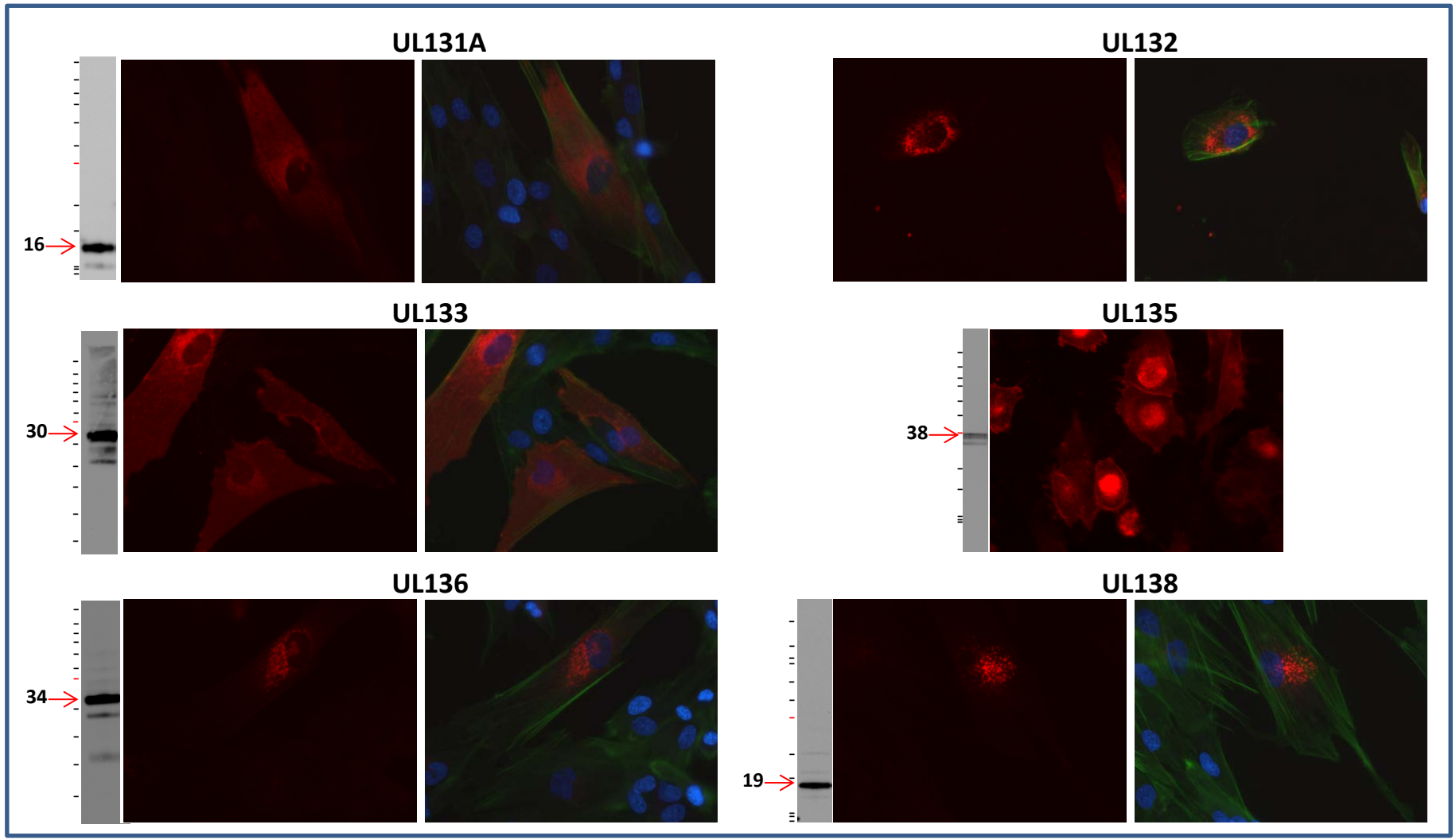

UL88
48
UL89
65
UL91
12
UL92
20
UL93
62
UL94
32

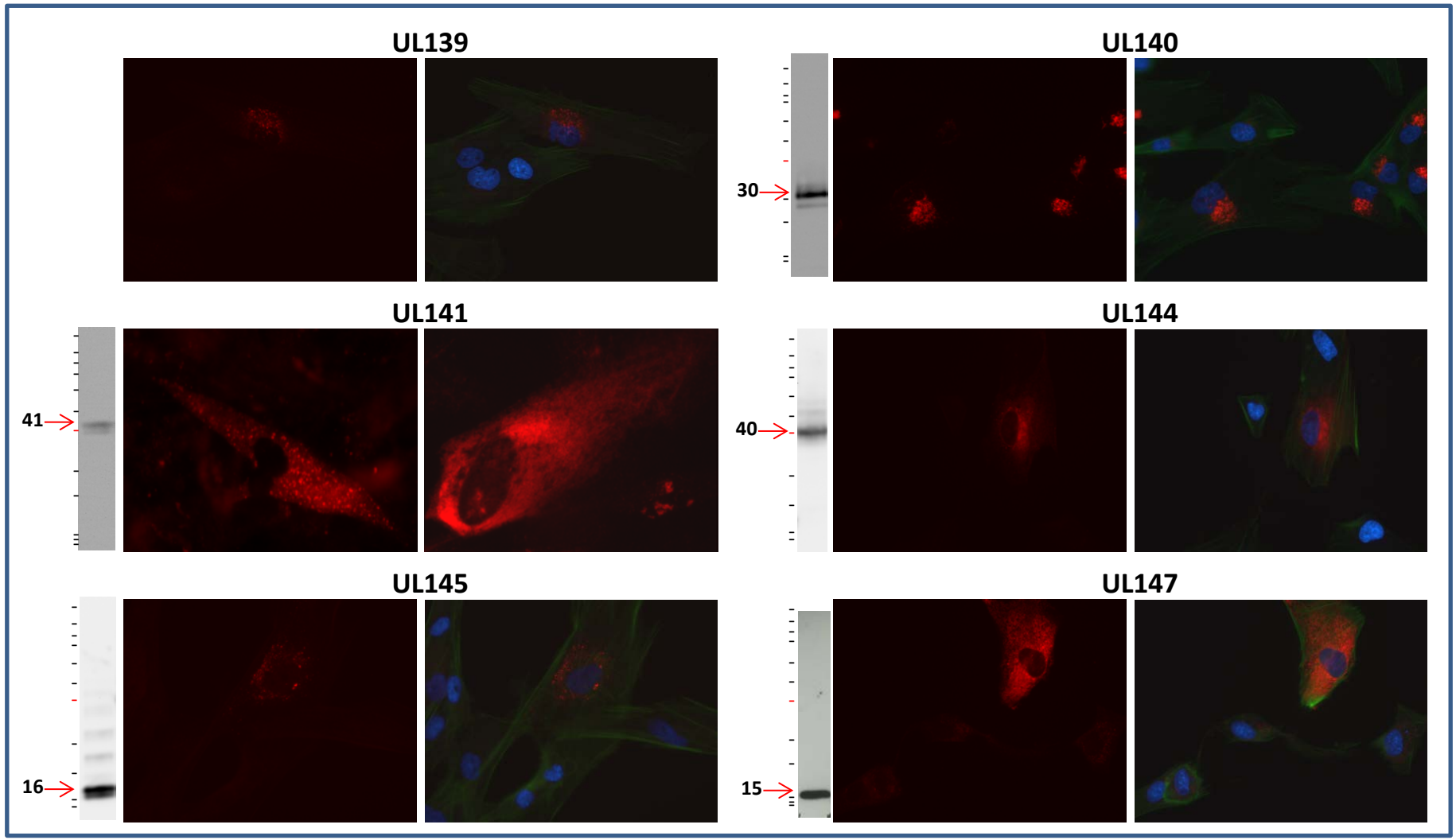

UL95
59
UL96
17
16
UL97
90
57
27
UL98
62
25
UL99
32
UL100

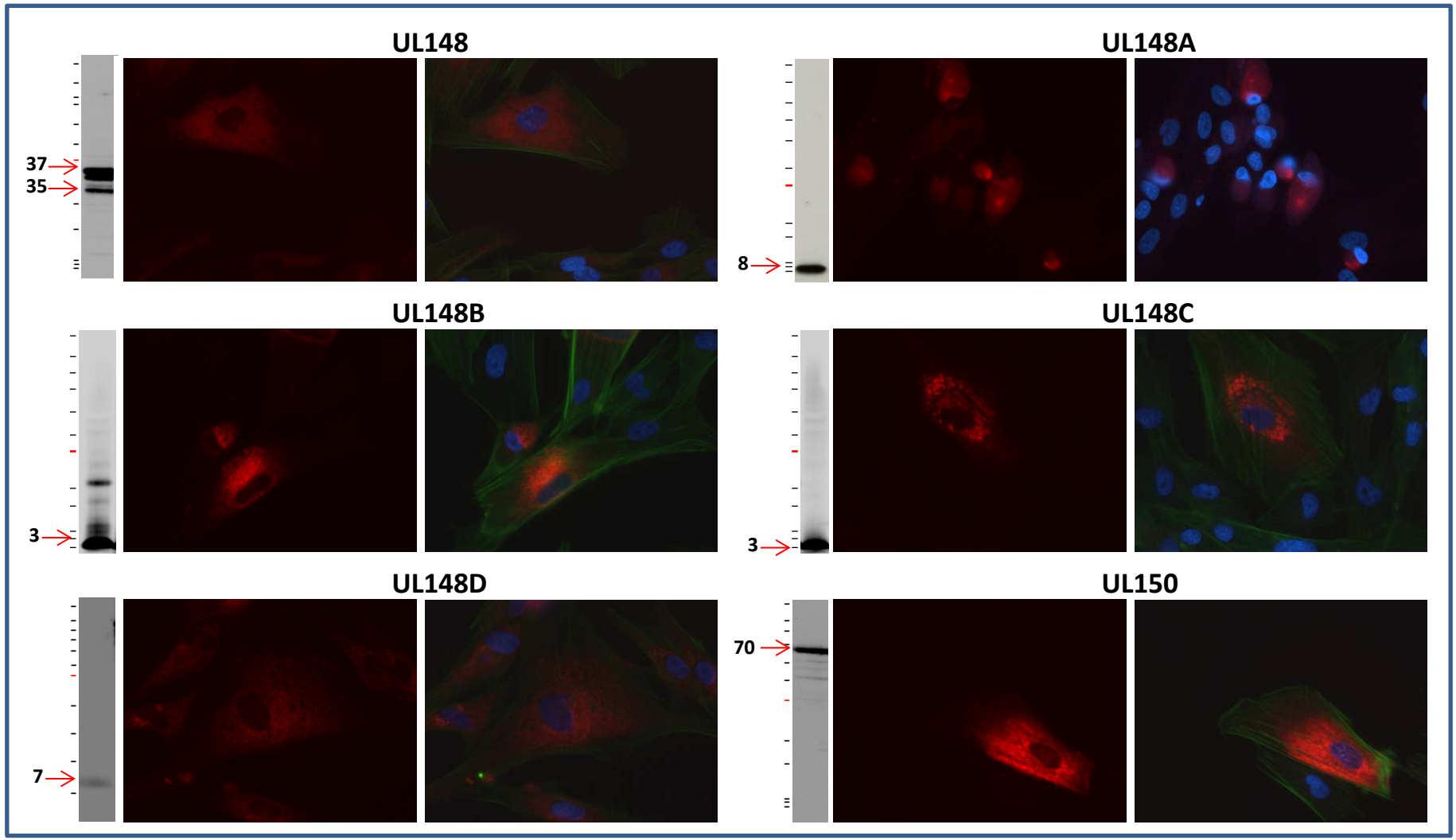

UL102
90
UL103
38
32
UL104
70
UL105
100
UL111A
19
17
UL112
82
42

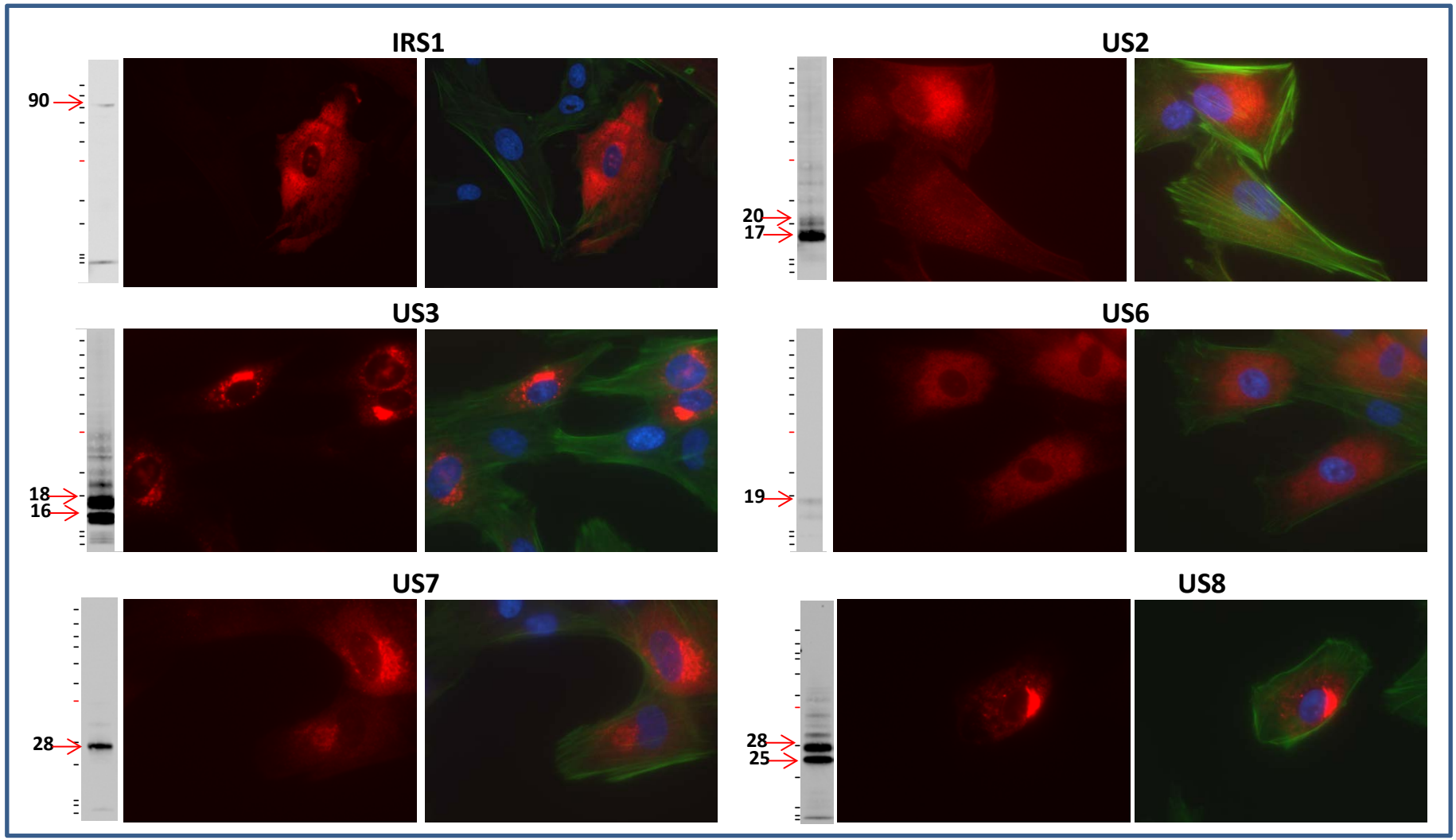

UL114
30
UL115
35
UL116
60
UL117
47
UL119
65
38
UL120
15

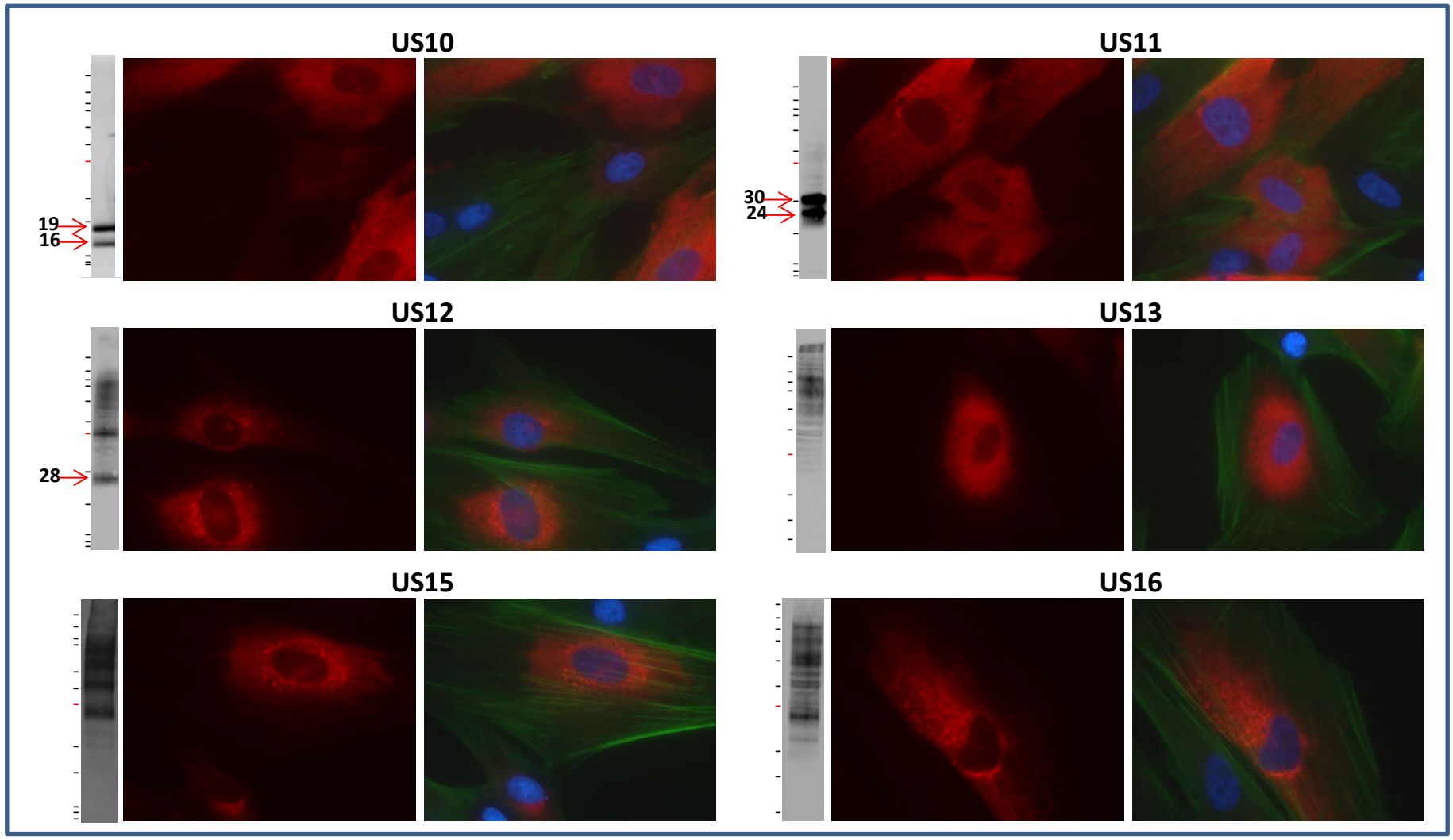

UL121
12
8
UL122
75
UL123
70
UL124
43
30
17
UL128
10
UL130
37

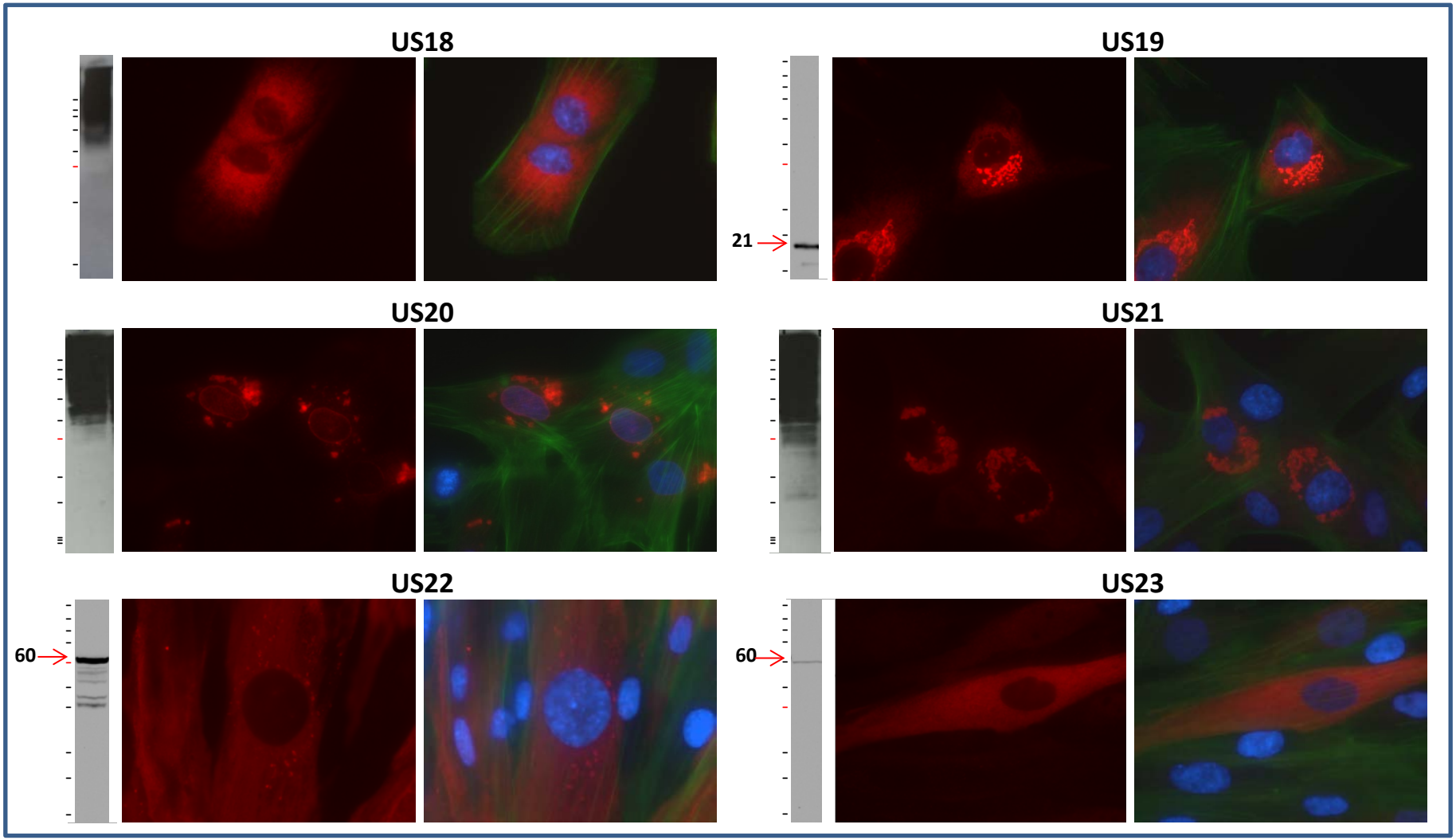

UL131A
16
UL132
UL133
30
UL135
38
UL136
34
UL138
19

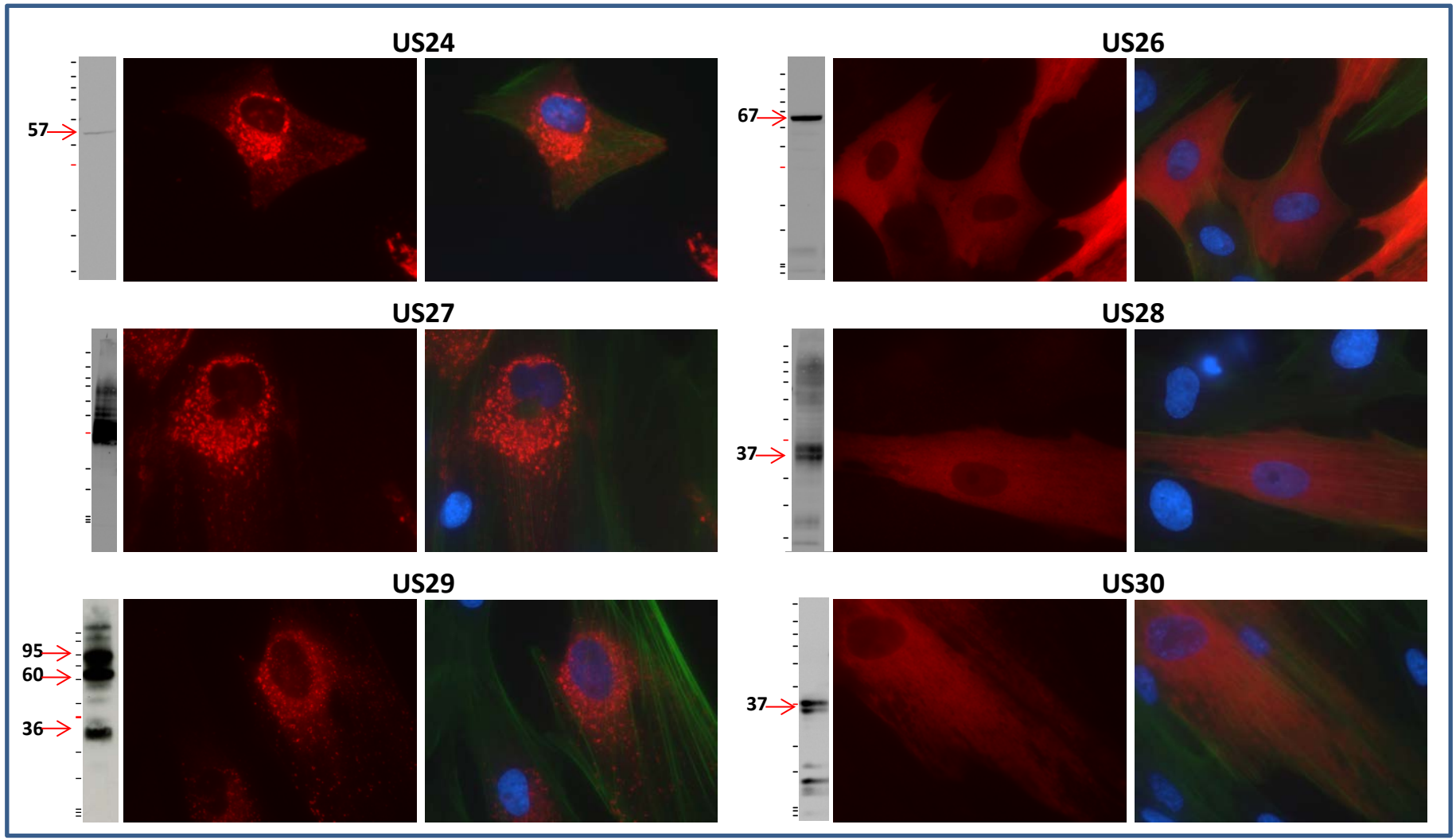

UL139
UL140
30
UL141
41
UL144
40
UL145
16
UL147
15


UL148
37
35
UL148A
8
UL148B
3
UL148C
3
UL148D
7
UL150
70


IRS1
90
US2
20
17
US3
18
16
US6
19
US7
28
US8
28
25


US10
19
16
US11
30
24
US12
28
US13
US15
US16


US18
US19
21
US20
US21
US22
US23
60
60


US24
57
US26
67
US27
US28
37
US29
95
60
36
US30
37

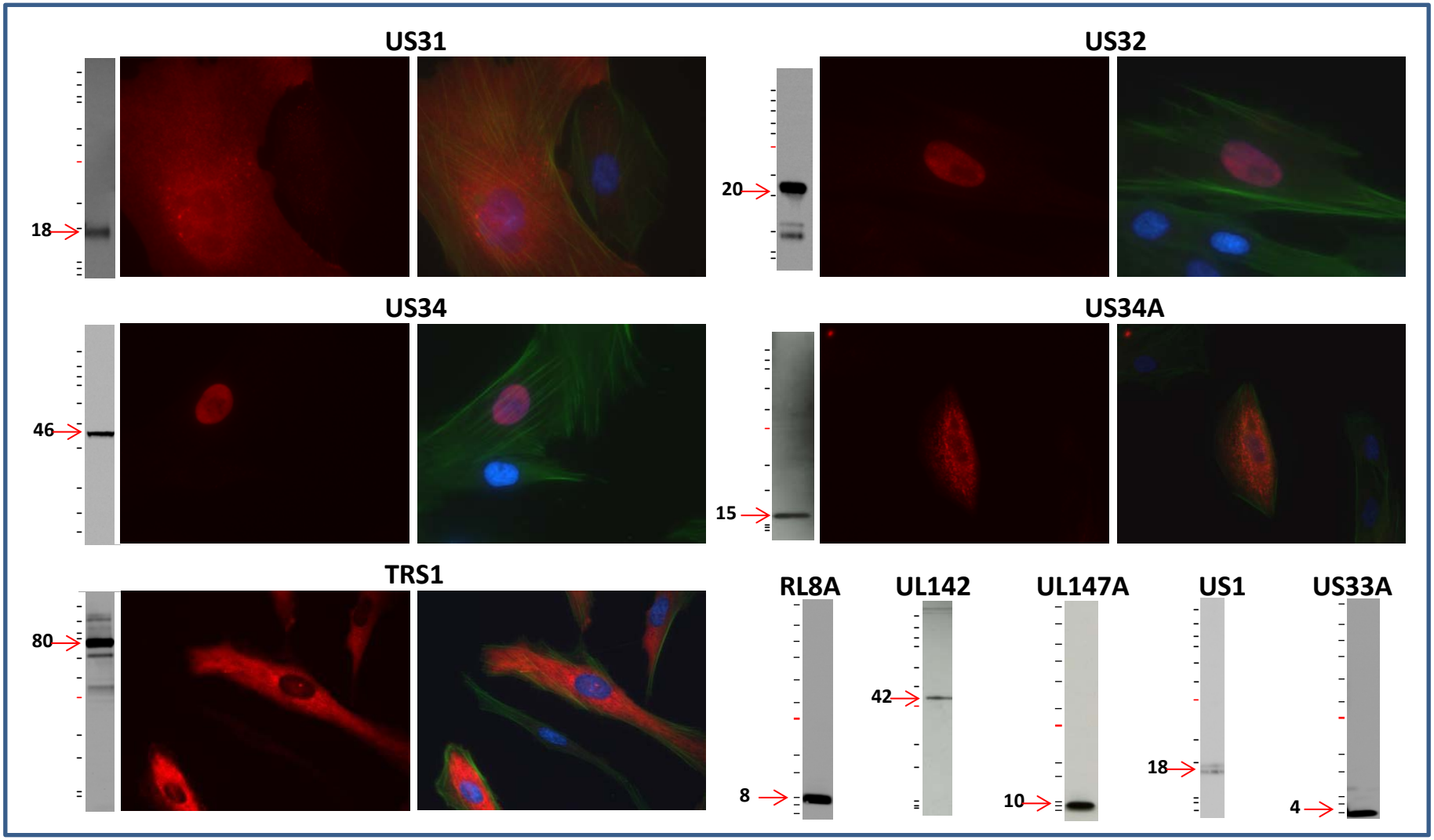

US31
18
US32
20
US34
46
US34A
15
TRS1
80
RL8A
8
UL142
42
UL147A
10
US1
18
US33A
4

Table 3.2: Description of data from western blot and immunofluorescence analyses of the HCMV RAd-library, and comparison with previous literature

From left to right, columns contain: gene name, predicted size of the primary translation product based on amino acid sequence (kDa), size of the predominant western blot species (kDa) (where more than one major species were observed sizes are given in order of abundance, unless specified otherwise. Very low molecular weight species adjudged to be break-down products were omitted), additional comments on the Western blot species detected, description of intracellular localisation observed, and previously published information on size and sub-cellular localisation. Where intracellular localisation observed was inconsistent with the previous literature, gene name is underlined. Gene name is in bold if the observed Western blot species varied by more than 10 kDa from previous report(s); where no previous reports were found, comparison was made with the predicted size. Neither reported nor predicted sizes are available for 16 uncharacterised glycoproteins. UL55 and UL74 genes were codon-optimised and are indicated with asterisks. Genes encoding known or putative glycoproteins (as predicted in Table 14.1, Davison and Bhella 2007) are shown in blue background, whilst genes that encode known or predicted phosphoproteins are shown in green background. UL32 encodes a protein that is both glycosylated and phosphorylated and is shown in blue background. Red background indicates expression was not detected. Please note expression of 8 ORFs was not detected by either western blot or immunofluorescence; these genes are exlcuded from this table.

Gene	Predicted polypeptide size (kDa)	predominant Western blot specie(s) (kDa)	Western blot: comments	Intra-cellular localisation by immunofluorescence	published information
RL1	35	38	Smaller and larger species also detected	Nuclear	Nuclear expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 37 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
RL5A	11	Not detected	N/A	Sub-cytoplasmic	N/A
RL8A	10	8	N/A	Not detected	N/A
RL10	19	29	Smaller species also detected	Sub-cytoplasmic/ perinuclear	22 kDa and 23.5 kDa proteins were detected in lysates of fibroblasts infected with strains AD169 and TB40E, and Cos-7 cells transfected with an RL10-expressing plasmid (Spaderna et al. 2002). Localisation was sub-cytoplasmic in transfected Cos-7 cells.
RL11	27	28, 35	N/A	Juxtanuclear	IgG Fc-receptor (gp34): 34 kDa protein expressed on the cell surface and accumulated in the perinuclear region of HCMV-infected cells (Furukawa et al. 1975, Keller et al. 1976, Atalay 2002).

RL12	47	80	N/A	cytoplasmic, juxtanuclear,	Sub-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).
RL13	30	80,55	Smaller species also detected	Sub-cytoplasmic	The RL13 gene was C-terminally V5-tagged within the Merlin BAC and recombinant virus generated (Stanton et al. 2010). Tagged gpRL13 was expressed as 100 kDa and 55 kDa isoforms, and localised to virion assembly compartments of infected fibroblasts.
UL1	25	Not detected	N/A	Sub-cytoplasmic	In AD169-infected cells, HA-tagged UL1 migrated at 55 kDa and co-localised with structural proteins at virion assembly compartments (Shikhagaie et al. 2012).
UL2	7	12	N/A	Sub-cytoplasmic	9 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL4	17	50	N/A	Sub-cytoplasmic	gp48: 48 kDa glycoprotein detected in lysates of fibroblasts infected with strain Towne (Chang et al. 1989).
UL5	19	16, 18, 23	N/A	Juxtanuclear	Pan-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 21 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL6	31	50	Larger and smaller species also detected	Perinuclear	N/A
UL7	24	53	Smaller species also detected	Sub-cytoplasmic, juxtanuclear	UL7, GFP-tagged at the C-termins was detected around nuclei of transfected cells using FITC filter, but at the cell surface using monoclonal UL7 antibody; C-terminus shed from surface. Expression not detected in HCMV-infected cells (Engel et al. 2011).
UL8	36	33	Many larger and smaller species also detected	Cytoplasmic, juxtanuclear, surface	N/A
UL9	27	11, 23, 38, 40	N/A	Sub-cytoplasmic, juxtanuclear	N/A
UL10	28	40	larger species also detected	Sub-cytoplasmic	N/A

UL11	31	48	N/A	Pan-cytoplasmic, surface	Expression detected on surface of AD169-infected cells by flow cytometry; intracellular localisation not investigated (Hitomi et al. 1997).
UL13	54	47	Smaller and larger species also detected	Pan-cytoplasmic, surface	N/A
UL14	36	40	N/A	Sub-cytoplasmic	In lysates of HCMV-infected fibroblasts, gpUL14 migrated at 35 kDa (Dan Cochrane, PhD thesis). Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).
UL15A	11	15	Smaller and larger species also detected	Sub-cytoplasmic	N/A
UL16	26	47	Smaller and larger species also detected	Juxtanuclear	50 kDa protein detected in AD169-infected cells (Kaye et al. 1992). UL16-EGFP co-localised with the ER-resident protein disulfide-isomerase (Welte et al. 2002).
UL17	13	16		Sub-cytoplasmic	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 15 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL18	42	72, 80	N/A	Pan-cytoplasmic	Cell surface MHC class-I homologue: Vaccinia virus- and RAd-expressed gpUL18 migrated at 67 kDa (Park et al. 2002). A larger 110 kDa isoform was identified during productive Merlin and AD169 infections; RAd-expressed gpUL18 accumulated intracellularly when detected by Immunofluorescence (Griffin et al. 2005).
UL19	11	14	N/A	Sub-cytoplasmic	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 13 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL20	38	38	Many larger and smaller species also detected	Sub-cytoplasmic	60 kDa and 80 kDa proteins detected in transfected 293 cells; intracellular gpUL20 was detected within endosomes and lysosomes (Jelcic et al. 2011).

UL21A	14	16	Doublet	Pan-cytoplasmic	15 kDa protein detected by immunoblotting AD169-infected fibroblast lysates and transfected HeLa cells; pUL21a was localised to the cytoplasm (Fehr and Yu, 2010).
UL22A	15	23	N/A	Sub-cytoplasmic	R27080/UL21.5/UL20a: major 45 kDa and minor 25 kDa glycoproteins detected in culture medium of AD169-infected fibroblasts; transient expression in Cos-7 cells produced 25 kDa secreted glycoprotein (Mullberg et al. 1999). Secreted gpUL22A from culture media of AD169-infected cells migrated at 27 kDa (Wang et al. 2004). UL22A mRNA packaged into virions was translated into Golgi-localised product at IE infection times. RAd-expressed gpUL22A-YFP and HCMV-derived gpUL22A also localised to the Golgi (Bresnahan and Shenk 2000).
UL23	33	28	Many larger and smaller species also detected	Pan-cytoplasmic, juxtanuclear, surface, sub-nuclear	33 kDa protein detected in purified AD169 virions but not infected cell lysates by Western blot. Juxtanuclear protein aggregates detected by immunofluorescence (Adair et al. 2002).
UL24	34	33	Many larger and smaller species also detected	Sub-cytoplasmic	40 kDa band detected in lysates of cells infected with AD169; pUL24 was located in a juxtanuclear cytoplasmic compartment (Adair et al. 2002).
UL25	74	73	Many smaller species also detected	Sub-cytoplasmic, Perinuclear	Putative phosphoprotein; several proteins detected between 80 kDa and 100 kDa in transfected and AD169-infected cells, including a predominant 85 kDa species (Battista et al. 1999). Immunofluorescence staining identified perinuclear and juxtanuclear cytoplasmic aggregates.
UL26	21	21	N/A	predominantly nuclear	Western blot performed on AD169-infected cell lysates identified major 21 kDa, and minor 26 kDa species (Stammninger 2002). In transfected cells, the 26 kDa species was more prominant. pUL26 localised to nuclei of infected and transfected cells.
UL27	69	Multiple, including 63	Many smaller species and one species >300 kDa also detected	Mostly nuclear	Locus of maribavir resistance: nuclear and nucleolar localisation observed for pUL27-GFP fusion protein in transfected cells by immunofluorescence (Chou et al. 2004).

UL28/29	79	68	some smaller species also detected	Sub-nuclear	Eptiope-tagged UL28/29 was detected by Western blot as a 79 kDa protein, and by immunofluorescence localised in nuclear viral replication centres (Mitchell et al. 2009).
UL30	14	19	Smaller species also detected	Pan-cytoplasmic, surface	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 16 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL31	66	68	some smaller bands also present	Sub-cytoplasmic, perinuclear	N/A
UL32	113	150	N/A	Pan-cytoplasmic	Major tegument phosphoglycoprotein (pp150): 150 kDa protein detected in fibroblasts infected with AD169 (Sanchez et al. 2000). Both nuclear (Hensel et al. 1995) and cytoplasmic (Sanchez et al. 1998) localisations reported during productive HCMV infection.
UL34	45	54	Doublet	Nuclear	43 kDa and 40 kDa proteins detected at early and late times of HCMV infection respectively; both proteins localised to the nucleus when expressed from plasmids (LaPierre and Biegalke 2001, Biegalke et al. 2004)
UL35	73	70	Many smaller species also detected	Nuclear, nuclear membrane, juxtanuclear	In fibroblasts infected with the Towne strain, UL35 encoded 22 kDa and 75 kDa proteins which localised to nuclei and juxtanuclear compartments repsectively (Liu and Biegalke 2002).
UL36	57	48	Many larger and smaller species also detected	Punctate cytoplasmic	Viral inhibitor of caspase activation (vICA): 52 kDa doublet of proteins detected in the lysates of cells infected with strains AD169, Towne and Toledo (Patterson and Shenk 1999). AD169-expressed pUL36 displayed uniform nuclear and cytoplasmic staining, whilst pUL36 expressed from Towne and Toledo was present in intense cytoplasmic foci.

UL37	56	65	Smaller species also detected	Sub-cytoplasmic	Viral mitochondira-localised inhibitor of apoptosis (vMIA): *in vitro*- synthesised UL37 primary protein and AD169-expressed gpUL37 migrated at 60 kDa and 83-85 kDa respectively (Al-Barazi and Colberg-Poley 1996). Transfected gpUL37 was detected as perinuclear granules, and co-localised with ER and Golgi proteins. Both trafficking through ER and Golgi apparatus to mitochondria (Colberg Poley et al. 2000), and complete mitochondrial localisation (Goldmacher et al. 1999) have been observed.
UL38	37	40	Doublet	Pan-cytoplasmic, surface	37 kDa pUL38 identified in the lysates of cells infected with AD169; shown by immunofluorescence to accumulate within nuclei, and to a lesser extent the cytoplasm of infected cells (Terhune et al. 2007).
UL40	24	20	Many larger and smaller species also detected	Sub-cytoplasmic	25 kDa and 28 kDa bands identified by immunoblotting in RAd-UL40-infected and AD169-infected cells (Tomasec et al. 2000). Pan-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).
UL41A	9	16	Smaller species also detected	Sub-cytoplasmic	N/A
UL42	14	18	Many larger and smaller species also detected	Sub-cytoplasmic, Perinuclear	N/A
UL43	48	39,50	Smaller species also detected	Pan-cytoplasmic, surface	48 kDa band detected in lysates of cells infected with AD169; pUL43 was located in juxtanuclear cytoplasmic compartments (Adair et al. 2002).
UL44	46	50	Many larger and smaller species also detected	nuclear	Processivity subunit of DNA polymerase/polymerase accessory protein (PAP/pp52): heterogeneously-sized phosphoprotein most abundantly migrating at 52 kDa, as detected in HCMV-infected cell lysates by Western blot (Weiland et al. 1994, Wolf et al. 2001). Immunofluresccence staining revealed punctate localisation for ppUL44 in nuclei of AD169-infected cells (Sourvinos et al. 2007). When DNA replication was inhibited in cells undergoing productive AD169 infection, ppUL44 was dispersed throughout the nucleus (Strang et al. 2012).

UL45	102	110, 80	Many larger and smaller species also present	Pan-cytoplasmic	Large subunit of ribonucleotide reductase: a 108 kDa protein was detected in lysates of fibroblasts undergoing productive AD169 infection, which co-migrated with pUL45 expressed in transfected HEK-293 cells (Patrone et al. 2003). Immunofluorescence analysis showed pUL45 is cytoplasmic during productive AD169 infection.
UL46	33	29	many larger and smaller proteins also present	Pan-cytoplasmic	Minor capsid protein-binding protein (mCP-BP): identified as a 30 kDa protein in AD169-infected fibroblasts (Gibson et al. 1996a).
UL47	110	110	Smaller species also detected	Nuclear	pUL47 localised to nuclei of fibroblasts infected with AD169, and was detected as a 110 kDa protein by Western blot (Bechtel and Shenk 2002).
UL48	253	250	N/A	Punctate sub-cytoplasmic	High molecular weight tegument protein: 216 kDa protein identified in the lysates of HCMV-infected fibroblasts; diffuse, as well as distinctive circular staining was observed in infected-cell cytoplasms (Bradshaw et al. 1994). pUL48 migrated at 250 kDa when extracted from fibroblasts infected with AD169 (Bechtel and Shenk 2002).
UL48A	8	8	N/A	Sub-nuclear, excluded from nucleoli	Small capsid protein (SCP): 9 kDa protein identified in lysates of fibroblasts infected with AD169 (Gibson et al. 1996). Nuclear and cytoplasmic staining observed when transiently expressed in isolation in Cos-7 cells, but accumulated in the cytoplasm when co-expressed with pUL86 (Lai and Britt 2003). Nuclear in AD169-infected fibroblasts.
UL49	64	55	N/A	Pan-cellular	N/A
UL50	43	44	Smaller species also detected	Nucler membrane	Inner nuclear membrane protein: Western blot analysis of epitope-tagged pUL50 transiently expressed in 293T cells identified a major 47 kDa and several faster migrating species (Milbradt et al. 2007) . Transiently expressed pUL50 localised to nuclear memranes and lamina (Milbradt et al. 2007, Miller et al. 2010).
UL51	17	22	N/A	Pan-cytoplasmic, surface	When tagged within AD169, pUL51 localised to the nuclei of infected fibroblasts and was detected as a single band by Western blot (size not provided) (Glab et al. 2009). 19 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).

UL52	74	75, 59	A smaller species also detected	Nuclear, sub-cytoplasmic	75 kDa protein detected by Western blot in the lysates of fibroblasts infected with recombinant AD169. Localisation was nuclear, surrounding the DNA replication and packaging compartments (Borst et al. 2008).
UL53	42	42	N/A	Nuclear	Nuclear matrix protein: Western blot analyses of transiently-expressed pUL53 identified a 45 kDa protein and some less abundant, faster migrating proteins (Milbradt et al. 2007); or, a 42 kDa protein (Dal Monte et al. 2002). Transiently expressed pUL53 was nuclear, but re-localised to the nuclear envelope when co-expressed with pUL50 (Milbradt et al. 2007, Miller et al. 2010). AD169-expressed pUL53 was weakly expressed in the cytoplasm at 48 hours p.i., but became juxtanuclear at 72 hours p.i. (Dal Monte et al. 2002).
UL54	137	140	Smaller species also detected	Nuclear	Catalytic subunit of DNA polymerase: *in vitro* translation resulted in expression of a 140 kDa protein (Ducancelle et al. 2006). Transfected Cos-7 cells expressed pUL54 in the nuclei (Alvisi et al. 2006).
UL55*	102	150, 55	Smaller species also detected	Sub-cytoplasmic	Glycoprotein B (gB): 160 kDa precursor protein cleaved to generate the mature gp55-gp116 complex (gCI) incorporated into virion envelopes (Britt and Auger 1986). When expressed in isolation gB is predominantly in the uncleaved form (Wells et al. 1990). gB localised to viral assembly compartments at juxtanuclear positions of AD169-infected cells (Womack and Shenk 2010). Localisation was perinuclear in transiently transfected Cos-7 cells and during productive HCMV infection (Krzyzaniak et al. 2007).
UL56	96	47, 110	Many larger and smaller species also detected	Nuclear	Large terminase subunit (p130): 130 kDa protein expressed by the UL56 gene was detected in AD169-infected fibroblasts (Bogner et al. 1998). Transiently-transfected pUL56 was nuclear but not nucleolar in Cos-7 and U373 cells (Giesen et al. 2000). In AD169-infected fibroblasts, pUL56 was (i) present in nuclei, including viral replication centres (Giesen et al. 2000), (ii) nuclei and particulary nucleoli at early times of infection, but became concentrated in discrete subnuclear structures and excluded from nucleoli at late times post-infection (Krzyzaniak et al 2007).

UL57	134	140	Many smaller species also detected	Punctate-nuclear	Single-stranded DNA-binding protein: 130 kDa protein detected in the lysates of fibroblasts infected with strains AD169 and Colburn (Anders et al. 1986). Fibroblasts infected with AD169 expressed nuclear pUL57; nuclear staining pattern was diffuse at 24 h p.i., but aggregated into focal globular structures at late times of infection (Kemble et al. 1987).
UL69	82	105	Smaller species also detected	Punctate-nuclear	pUL69 localised to the nuclei of fibroblasts infected with AD169 and migrated at 105 kDa (Bechtel and Shenk 2002).
UL70	108	110	Many larger and smaller species also detected	Pan-cytoplasmic	Component of DNA-helicase complex: Tagged UL70 migrated at ~100 kDa, and was predominantly cytoplasmic in transfected UL373MG cells. However, tagged pUL70 was efficiently imported into the nucleus when these cells were infected with strain Towne (Shen et al. 2011).
UL71	40	49	Some larger species also detected	Juxtanuclear	GFP-tagged pUL71 localised to viral assembly compartments in fibroblasts infected with AD169 (Womack and Shenk 2010). pUL71 was detected in lysates of fibroblasts infected with strain TB40, and transiently transfected Cos-7 cells at 48 kDa (Schauflinger et al. 2011).
UL72	43	42	N/A	Sub-cytoplasmic, some surface	A 46 kDa protein was detected by Western blot in lysates of fibroblasts infected with AD169, or transiently expressing the UL72 gene from a plasmid (Caposio et al. 2004). In both cases, pUL72 showed cytoplasmic immunofluorescence staining.
UL73	14	17, 34	Intermediately-sized species also detected	Juxtanuclear	Glycoprotein N (gN): pUL73 expressed in transiently transfected 293T cells migrated at 18 kDa, but was post-tranlationally modified when co-expressed with gM (UL100) to form a 50-60 kDa product (Mach et al. 2000). The localisation of pUL73 in Cos-7 cells was limited to the ER and ERGIC when expressed alone, but extended to the cellular secretory pathway in the presence of gM. AD169-expressed gM/gN complex was detected within juxtanuclear assembly compartments of infected Cos-7 cells (Krzyzaniak et al. 2007).

UL74*	54	85	Smaller species also detected	Pan-cytoplasmic	Glycoprotein O (gO): 100-125 kDa glycoprotein in the context of HCMV infection (size is glycosylation-state and strain-dependent); forms a complex with gH/gL (220-260 kDa) which transits through the secretory pathway (Huber and Compton 1998, Huber and Compton 1999). localised to juxtanuclear compartments in AD169-infected fibroblasts (Theiler and Compton 2002). Isolated expression resulted in faster migration in gel and incorrect trafficking.
UL75	84	80	N/A	Sub-cytoplasmic, nuclear membrane	Glycoprotein H (gH): in the context of HCMV infection, forms complexes with gL/gO (220-260 kDa), or gL/UL128/UL130/UL131A, which transit through the secretory pathway (Huber and Compton 1999). Detected as an 86 kDa protein in fibroblasts infected with AD169, or recombinant vaccinia virus expressing the UL75 gene (Cranage et al. 1988). AD169-expressed gH was present at the nuclear membrane, cell surface, or in cytoplasmic bodies (now known to be assembly compartments). In contrast, gH was present predominantly at the nuclear membrane when expressed alone. Co-expression with gL relieves ER retention and allows cell surafce expression (Spaete et al. 1993).
UL76	36	Not detected	N/A	Sub-nuclear	In COS-1 cells transiently expressing UL76, and AD169-infected fibroblasts, 38 kDa pUL77 was detected by Western blot. GFP-tagged pUL76 aggregated in sub-nuclear foci of transfected COS-1 cells and fibroblasts (Wang et al. 2000).
UL77	71	Not detected	N/A	Sub-cytoplasmic	When tagged within AD169, pUL77 localised nuclei of infected fibroblasts (Glab et al. 2009). 73 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL78	47	40	Some larger species also detected	Sub-cytoplasmic	When transiently expressed in HeLa and 293T cells, or during infection of fibroblasts with TB40-BAC4-derived-recombinant HCMV; pUL78 localised to the cytoplasm and associated mainly with the ER and early endosomes (Wagner et al. 2012). A 45 kDa protein was detected by Western blot in both settings.

UL79	34	32	Doublet; smaller species also detected	Nuclear/ cytoplasmic	HA-tagged pUL79 migrated at 35 kDa and localised to nuceli and nuclear replication compartments when stably expressed in fibroblasts. When these cells were infected with recombinant AD169 encoding conditionally expressed pUL79, or harbouring a UL79 gene deletion, HCMV-expressed-pUL79 could not be detected (Perng et al. 2011). pUL79 epitope-tagged in strain Towne localised to replication compartments during productive infection of fibroblasts, and migrated at about 35 kDa (Isomura et al. 2011).
UL80	74	78	N/A	Nuclear	Protease/minor capsid scaffold protein: expressed in the nuceli of cells transfected with GFP-tagged pUL80, and fibroblasts infected with AD169 encoding epitope-tagged UL80 (Nguyen et al. 2008). Synthesised as a 74 kDa protein (Plafker and Gibson 1998).
UL80.5	38	50	N/A	Punctate nuclear	Assembly protein precursor/major capsid scaffold protein, present in NIEPs but not virions (pAP): 40 kDa protein translated *in vitro* from cDNA (Roby and Gibson 1986). Other studies detected 38 kDa protein; pUL80.5 is present throughout the nucleus of transfected Cos-7 cells (Wood et al. 1997, Plafker and Gibson 1998).
UL82	62	60	Many smaller species also detected	Punctate nuclear	Upper matrix protein (pp71): migrated at 71 kDa when lysates of AD169-infected cells were analysed by Western blot (Hensel et al. 1996). Localised to nuclei of fibroblasts infected with AD169 (Hensel et al. 1996, Bechtel and Shenk 2001). Transiently expressed pp71 localised to discrete dot-like nuclear domains in astrocytoma cells (Hensel et al. 1996).
UL83	63	60	Many larger and smaller species also detected	Nuclear	Lower matrix protein (pp65): localised to the nuclei of fibroblasts infected with AD169, and migrated at 65 kDa (Sanchez et al. 1998, Bechtel and Shenk 2001). Transiently-expressed pp65 was evenly distributed in astrocytoma cell nuceli (Hensel et al. 1996).
UL84	65	68	Many smaller species also detected	Nuclear	65 kDa protein detected in the lysates of fibroblasts infected with the Towne strain (He et al. 1992). Localised to the nucleus and the cytoplam.
UL85	35	34	Many larger and smaller species also detected	Pan-cytoplasmic, surface	Minor capsid protein (mCP): expressed predominantly in the nuclei of AD169-infected fibroblasts (Adamo et al. 2004). 34 kDa protein incorporated into HCMV virions (Roby and Gibson 1986).

UL86	154	150	N/A	Pan-cytoplasmic, surface	Major capsid protein (MCP): pUL86 migrated at 150 kDa when extracted from fibroblasts infected with AD169 (Bechtel and Shenk 2001). Perinuclear/cytoplasmic localisation when transiently expressed in Cos-7 cells but nuclear in the context of AD169 infection of fibroblasts (Sanchez et al. 1998, Lai and britt 2003). Nuclear translocation requires presence of pUL80.5 (Wood et al. 1997).
UL87	105	90	N/A	Pan-cellular/ sub-nuclear	The UL87 gene was epitope-tagged within the Towne strain; Western blot analysis of infected cell lysates revealed proteins migrating between 100 kDa and 150 kDa (Isomura et al. 2011). pUL87 was localised to replication compartements within nuclei of infected fibroblasts.
UL88	48	48	Many larger and smaller species also detected	Pan-cellular	Coomassie staining of purified AD169 particles revealed a 48 kDa virion component. Protein sequencing showed this was encoded by the UL88 gene (Baldick and Shenk 1996).
UL89	77	65	N/A	Pan-cytoplasmic, surface	Small terminase subunit: Western blot analysis of AD169-infected HFF lysates detected a 75 kDa protein. Immunofluorescence staining showed pUL89 localised to replication compartments within nuclei of infected cells (Thoma et al. 2006).
UL91	12	12	N/A	Sub-cytoplasmic	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 14 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL92	23	20	Many larger and smaller species also detected	Pan-cellular	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 25 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL93	68	62	N/A	Pan-cytoplasmic/ nuclear	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 71 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).

UL94	38	32	Larger and smaller species also detected	Nuclear	pUL94 detected as a 36 kDa protein by *in vitro* transcription and translation of the UL94 gene in lysates of 293T cells stably expressing the UL94 gene, and in lysates of fibroblasts infected with the Towne strain (Wing et al. 1996). Transiently expressed pUL94 localised to nuclei of Vero cells, but was both nuclear and cytoplasmic when co-expressed with pp28. pp28-mediated traslocation of pUL94 to juxtanuclear assembly compartments during productive HCMV infection of fibroblasts was also observed (Liu et al. 2009).
UL95	57	59	Doublet	Pan-cellular	59 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL96	14	16, 17	N/A	Sub-cytoplasmic	15 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL97	78	90, 57, 27	Many other species also detected	Nuclear	Serine-threonine protein kinase: 80 kDa protein, and some less abundant faster migrating species were detected in the lysates of HCMV-infected fibroblasts by Western blot (Littler et al. 1992). pUL97 localised to the nuclei of AD169-infected fibroblasts (Michel et al. 1996). Transient pUL97 expression using a plasmid was detected in the nuclei of Cos-7 cells (Prichard et al. 2005a).
UL98	65	25, 62	Many other species also detected	Pan-cytoplasmic, perinuclear, surface	Deoxyribonuclease: 58 kDa protien detected by Western blot in lysates of fibroblasts infected with AD169 (Adam et al. 1995). Immunofluorescence staining of fibroblasts infected with the Towne and CR208 strains identified pUL98 to be nuclear (Gawn and Greaves 2002).
UL99	21	32	Larger and smaller species also detected	Juxtanuclear	Myristylated tegument phosphoprotein (pp28): 28 kDa product detected in lysates of fibroblasts infected with the Towne strain (Re et al. 1985). Localised to cytoplasmic assembly compartments in fibroblasts during the course of AD169 infection (Sanchez et al. 2000).

UL100	43	Not detected	N/A	Sub-cytoplasmic	gM: 45 kDa protein complexed with gN through disulfide linkage (Mach et al. 2000). Localised to the ER and ERGIC in Cos-7 cells transiently-transfected with a UL100 expression vector but was re-localised to compartments of the secretory pathway when co-expressed with gN. AD169-expressed gM/gN complex was detected within juxtanuclear assembly compartments of infected Cos 7 cells (Krzyzaniak et al. 2007).
UL102	94	90	N/A	Pan-cytoplasmic, surface	N/A
UL103	29	32, 38	N/A	Sub-cytoplasmic, Perinuclear	Localised to juxtanuclear assembly compartments in cytoplasm of fibroblasts infected with the Towne strain (Ahlqvist and Mocarski 2011). 31 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL104	79	70	some smaller species also detected	Pan-cellular, surface	Portal protein: pUL104 localised to nuclei of fibroblasts infected with AD169; diffusely cytoplasmic at early times of infection, but at late times of infection aggregated into ring-shaped structures thought to be sites of DNA packaging, as well as punctate perinuclear cytoplasmic structures (Dittmer et al. 2005). The pUL104 protein migrated at 73 kDa.
UL105	106	100	some smaller species also detected	Pan-cellular, surface	Helicase: migrated at 105 kDa by Western blotting of AD169-infected fibroblast, and when transiently expressed (Smith et al. 1996).
UL111A	20	17, 19	N/A	Cytoplasmic	CMV-IL10: Cos-7 cells were transiently transfected with a plasmid encoding epitope-tagged UL111A; a 21 kDa product was detected by Western blot in conditioned cell culture media. Minor 30-35 kDa glycosylated species were also detected that could be reduced to 21 kDa by PNGaseF treatment (Kotenko et al. 2000). Sub-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).

UL112	70	42, 82	Other species also present	Nuclear, punctate nuclear	During productive infection of fibroblasts with AD169, the UL112 gene encodes 34, 43, 50 and 84 kDa proteins with common N-termini, translated from a differentially spliced mRNA precursor (Wright and Spector 1989). The 43 kDa isoform was the most abundant of the four species. pUL112 was detected by immunofluorescene in viral replication compartemnts, and on the periphery of PML bodies (Yamamoto et al. 1998, Ahn et al. 1999).
UL114	28	30	N/A	Pan-cytoplasmic, surface	Uracil DNA glycosylase: Plasmid-expressed UL114 was cytoplasmic, but recruited to the nucleus when co-expressed with UL44 (Prichard et al. 2005). pUL114 tagged within AD169 was expressed in infected cell nuclei, and detected by western blotting as a 34 kDa product.
UL115	31	35	N/A	Sub-cytoplasmic	Glycoprotein L (gL): 32 kDa glycoprotein (Spaete et al. 1993). In the context of HCMV infection, forms a trimeric complex with gH/gO (220-260 kDa), or a pentomeric complex with gH/UL128/UL130/UL131A; these complexes transit through the secretory pathway (Huber and Compton 1999). Localised to juxtanuclear compartments in AD169-infected fibroblasts (Theiler and Compton 2002)
UL116	34	60	N/A	Sub-cytoplasmic	N/A
UL117	46	47	N/A	Sub-nuclear, punctate nuclear	In the context of HCMV infection, the UL117 gene produces 35 kDa (pUL117.5) and 45 kDa (pUL117) proteins which localise to the nuclei of infected cells (Qian et al. 2008). Nuclear pUL117 associates with viral replication compartments.
UL119	39	38, 65	N/A	Sub-cytoplasmic	IgG Fc-binding glycoprotein (gp68): 68 kDa protein, present in the cytoplasm, from where passes through endocytic pathway to the surface of HCMV-infected cells; Immunofluorescence detecion not performed (Stannard and Hardie 1991, Atalay 2002).
UL120	23	15	N/A	Pan-cytoplasmic, surface	N/A

UL121	20	8, 12	Smaller species also detected	Sub-cytoplasmic	Pan-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).
UL122	63	75	N/A	Nuclear	IE2: encodes multiple spliced gene products including a major 86 kDa protein (IE 86), localises to the nucleus (Wilkinson et al. 1985). 82-86 kDa species routinely detected by Western blot (Harel and Alwine 1998).
UL123	55	70	N/A	Nuclear	IE1: encodes a 72 kDa phosphoprotein (IE72) that localises to the nucleus (Wilkinson et al. 1998).
UL124	17	17, 30, 40	N/A	Sub-cytoplasmic	Sub-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).
UL128	20	10	N/A	Sub-cytoplasmic	The 16 kDa pUL128 forms a complex with gH/gL/UL130/UL131A, which is incoporated into virion envelopes (Wang and Shenk 2005). pUL128 expressed from a plasmid was detected in the cytoplasm of 293T cells (Gerna et al. 2008).
UL130	25	37	some smaller species also detected	Sub-cytoplasmic	The UL130 gene produces a 33/35 kDa doublet of proteins (Wang and Shenk 2005). pUL130 forms complexes with gH/gL/UL128/UL131A, which are incorporated into virion envelopes. UL130 expressed from a plasmid was detected in the cytoplasm and on the surface of 293T cells (Gerna et al. 2008).
UL131A	15	16	Smaller species also detected	Pan-cytoplasmic, surface	Associates with gH/gL/UL128/UL130 to form a high molecular weight complex (Wang and Shenk 2005). 17 kDa pUL131A detected in the lysates of fibroblasts infected with TB40. pUL131A expressed from a plasmid was detected in the cytoplasm of 293T cells (Gerna et al. 2008).
UL132	30	Not detected	N/A	Juxtanuclear	Two isoforms were detected in AD169-infected fibroblast: a 22-28 kDa species and a 45-60 kDa species. Immunofluorescence analysis showed pUL132 co-localised with markers of the Golgi apparatus and the trans-Golgi network (Spaderna et al. 2005).

UL133	28	30	Larger and smaller species also detected	Pan-cytoplasmic, surface	A product migrating near to the 35 kDa marker was detected in TB40E-infected fibroblast lysates by Western blot (Petrucelli et al. 2012). A 28 kDa (excluding the size of the tag) tagged UL133 product was detected by immunoblotting of transiently transfected HEK-293 cell lysates (Grainger et al. 2010). When UL133 was tagged within he FIX virus, a protein of similar size was detected in infected fibroblasts.
UL135	33	38	Doublet	Sub-cytoplasmic, some surface	37 kDa doublet detected by Western blot in lysates of fibroblasts infected with strain Merlin (Dr Richard Stanton, unpublished data). Localised to Golgi and cell membrane
UL136	27	34	some smaller species also detected	Juxtanuclear	Excluding the size of its epitope tag, a 27 kDa UL136 protein was detected in the lysates of transiently transfected 293 cells (Grainger et al. 2010). In fibroblasts infected with a recombinant FIX virus encoding an apitope-tagged UL136, a protein of identical size was detected, in addition to a smaller protein.
UL138	19	19	N/A	Sub-cytoplasmic	Infection of fibroblasts with the FIX strain resulted in expression of a 21 kDa protein that was detected by Western blot using anti-UL138 serum (Petrucelli et al. 2009). In this context, pUL138 localised to the Golgi of infected cells.
UL139	14	Not detected	N/A	Juxtanuclear	N/A
UL140	21	30	Doublet	Juxtanuclear	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).
UL141	39	41	Doublet	Cytoplasmic	gpUL141 was detected by Western blot as a 37-40 kDa glycoprotein doublet in the lysates of fibroblasts infected with strains Toledo and AD169 (Tomasec et al. 2005). gpUL141 localised to the cytoplasm of AD169-infected fibroblasts, and transfected 293T cells.

UL142	35	42	N/A	Not detected	GFP tagged-UL142 transiently transfected into fibroblasts migrated at just over 67 kDa, and localisd to the cytoplasm; unpermeabilised cells expressed the fusion protein on the surface (Wills et al. 2005). Localisation of a transfected GFP-UL142 protein to the ER and *cis* -Golgi apparatus was observed by immunofluorescence in both un-infected, and TB40E-infected fibroblasts (Ashiru et al. 2009).
UL144	20	40	N/A	Juxtanuclear	Western blot analysis of UL144 gene products in fibroblasts infected with strain Fiala revealed a major protein of ~44 kDa, and a minor protein of 38 kDa (Benedict et al. 1999). pUL144 could not be detected by flow cytometry, and is likely retained intracellularly.
UL145	15	16	Some larger species also detected	Sub-cytoplasmic	Pan-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 13 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
UL147	19	15	N/A	Pan-cytoplasmic, some surface	N/A
UL147A	8	10	N/A	Not detected	N/A
UL148	35	35, 37	37 kDa band a doublet	Sub-cytoplasmic	N/A
UL148A	9	Not detected	N/A	Sub-cytoplasmic	N/A
UL148B	9	3	N/A	Sub-cytoplasmic	N/A
UL148C	9	3	Larger species up to 250 kDa also detected	Sub-cytoplasmic	N/A
UL148D	7	7	N/A	Sub-cytoplasmic	N/A
UL150	70	70	N/A	Sub-cytoplasmic	Pan-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 73 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).

IRS1	91	90	5 kDa species also detected	Cytoplasmic, sub-nuclear	Molecular mass of pIRS1 detected by Western blot in lysates of AD169-infected fibroblasts was consistent with the predicted size of 91 kDa (Romanowski and Shenk 1997). A smaller 30 kDa protein (designated pIRS1^{263}) was also recognised, thought to be the product of a 263-aa polypeptide utilising an alternative translation start site. In transfected HeLa cells, pIRS1 and pIRS1^{263} were predominantly cytoplasmic and nuclear respectively.
US1	18	18	N/A	Not detected	N/A
US2	23	17, 20	Some larger species also detected	Sub-cytoplasmic	24 kDa glycoprotein, and less abundant 21 kDa and 14 kDa species detected in the lysates of fibroblasts infected with strain AD169 (Jones and Sun 1997). Tagged gpUS2 expressed from a transfected expression vector was ER-localised in un-infected and HCMV Towne-infected extravillus trophoblasts and fibroblasts respectively (Terauchi et al. 2003).
US3	22	16, 18	Some larger species also detected	Sub-cytoplasmic, perinuclear, juxtanuclear	Transiently expressed gpUS3 migrated at 22 kDa in Hela cells (Ahn et al. 1997). 22 kD, 17 kDa and 3.5 kDa species detected in lysates of transfected HeLa cells (Shin et al. 2006). gpUS3 was ER-localised in HeLa cells when expressed in isolation , and in the context of HCMV-Towne infection (Ahn et al. 1996). gpUS3 expressed from a transfected expression vector was ER-localised in uninfected and HCMV Towne-infected extravillus trophoblasts and fibroblasts (Terauchi et al. 2003).
US6	21	19	N/A	Sub-cytoplasmic	Transiently expressed gpUS6 migrated at 22 kDa in Pala cells (Lehner et al. 1997). Transiently expressed US6 localised to the ER in HeLa cells (Ahn et al. 1997). Transiently expressed (HeLa cells), and AD169-derived (fibroblasts) US6 were ER-localised and migrated at 21 kDa (Hengel et al. 1997). Tagged US6 expressed from a transfected expression vector was ER-localised in uninfected and HCMV Towne-infected extravillus trophoblasts and fibroblasts (Terauchi et al. 2003).

US7	26	28	N/A	Sub-cytoplasmic	Immunofluorescence staining of RAd-expressed gpUS7 showed it was diffusely cytoplasmic, and co-localised with ER-resident proteins (Huber et al. 2002). In these cells, gpUS7 migrated closely with the 30 kDa protein marker.
US8	27	25, 28	Larger and smaller species also detected	Juxtanuclear	Transiently expressed gpUS8 was detected by Western blot as a 32-30 kDa doublet, and localised to ER and Golgi compartments in COS-1 and epithelial cells respectively (Maidji et al. 1998). When expressed from a RAd in HEC-1A cells, US8 was detected in punctate perinuclear vesicles and showed extensive co-localisation with markers of the Golgi apparatus (Huber et al. 2002).
US10	21	16, 19	N/A	Pan-cytoplasmic, some surface	Immunofluorescence staining of RAd-expressed gpUS10 showed it was diffusely cytoplasmic, and co-localised with ER-resident proteins (Huber et al. 2002). In these cells, gpUL10 migrated slightly faster than the 30 kDa markers, but exact size was not provided. HA-tagged UL10 protein migrated at 23 kD and was localised to the ER, in stably transfected U373 cells (Furman et al. 2002).
US11	25	24, 30	N/A	Sub-cytoplasmic	US11 expressed in AD169-infected fibroblasts migrated at 31 kDa (Jones et al. 1991). US11 was ER-localised in HeLa cells when expressed in isolation, and in the context of HCMV-Towne infection (Ahn et al. 1996). Tagged US11 expressed from a transfected expression vector was ER-localised in uninfected and HCMV Towne-infected extravillus trophoblasts and fibroblasts (Terauchi et al. 2003).
US12	32	28, 40->300	Includes continuous smear	Sub-cytoplasmic	N/A
US13	29	35->300	Includes continuous smear	Sub-cytoplasmic	N/A
US15	29	35->300	Includes continuous smear	Sub-cytoplasmic	55 kDa product detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).

US16	35	35->300	Includes continuous smear	Sub-cytoplasmic	The US16 gene was HA-tagged within the TR BAC and recombinant virus generated; Western blot analysis of infected fibroblast lysates showed pUS16 migrated at 33 kDa (Bronzini et al. 2012). pUS16 co-localised with gB and pp28 in the cytoplasm of infected cells, suggesting it is present in viral assembly compartments. A similar cytoplasmic pattern was observed in fibroblasts transiently transfected with a pUS16 expression vector.
US18	30	50->300	Includes continuous smear	Sub-cytoplasmic	36 kDa pUS18 detected in lysates of fibroblasts infected with the Towne strain by Western blot (Guo and Huang et al. 1993). pUS18 was detected by immunofluorescence in the cytoplasm of fibroblasts infected with AD169, concentrating at assembly compartments (Das et al. 2006, Das and Pellett 2007).
US19	26	21	N/A	Sub-cytoplasmic, punctate justanuclear	32 kDa pUS19 detected in lysates of fibroblasts infected with the Towne strain by Western blot (Guo and Huang et al. 1993). pUS19 was also detected by immunofluorescence in these cells but the authors did not show this data.
US20	29	40->300	Includes continuous smear	Sub-cytoplasmic, nuclear membrane	43 kDa pUS20 detected in the lysates of fibroblasts infected with the Towne strain by Western blot (Guo and Huang et al. 1992). pUS20 was also detected by immunofluorescence in these cells but the authors did not show this data.
US21	27	20->300	Includes continuous smear	Sub-cytoplasmic	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 29 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
US22	65	60	some smaller species also detected	Pan-cytoplasmic, some surface	Immunoblotting identified a 66 kDa protein in lysates of AD169-infected cells. pUS22 was found to be nuclear by immunofluorescence, but partitioned with the cytoplasm in fractionation studies (Mocarski et al. 1988).
US23	69	60	N/A	Pan-cytoplasmic, some surface	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). 71 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).

US24	58	57	N/A	Sub-cytoplasmic, perinuclear	Antibody specfic to pUS24 detected a 58 kDa protein in the lysates of fibroblasts infected with strain AD169. In these cells, Immunofluorescently stained pUS24 co-localised with pp65 in the cytoplasm (Feng et al. 2006).
US26	70	67	N/A	Pan-cytoplasmic, surface	Pan-cellular expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).
US27	42	40->300	N/A	Sub-cytoplasmic, perinuclear	Found at the cell surface, and more prominantly in association with perinuclear endocytic and lysosomal organelles both in transfected cells, and cells undergoing productive HCMV infection (Fraile-Ramos et al. 2002). 44 kDa protein detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
US28	41	37	Doublet, Many larger and smaller species also detected	Pan-cytoplasmic, surface	pUS28 detected in lysates of AD169-infected fibroblast, and transfected HEK293A cells migrated at 36 kDa (Mokros et al. 2002). Tagged pUS28 co-localised with markers of early/late endosomes and lysosomes in transfected HeLa cells (Fraile-Ramos et al. 2001). A pUS28-YFP fusion protein was intracellular in transiently transfected fibroblasts even after infection with the Towne strain. pUS28 is present at low levels on the surface of HCMV-infected (Michelson et al. 1997, Fraile-Ramos et al. 2001) and transfected cells (Pleskoff et al. 1998), but is rapidly endocytosed.
US29	51	36, 60, 95	N/A	Sub-cytoplasmic, perinuclear	N/A
US30	39	37	Doublet, other smaller species also detected	Sub-cytoplasmic	Sub-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).
US31	19	18	N/A	Pan-cytoplasmic, surface	Sub-cytoplasmic expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008).

US32	22	20	Smaller species also detected	Sub-nuclear	Sub-nuclear expression was observed by immunofluorescence staining of 293T cells transiently transfected with an expression plasmid (Salsman et al. 2008). A protein migrating at ~25 kDa was observd by Western blotting after adjusting for the 10 kDa SPA tag.
US33A	7	4	N/A	Not detected	N/A
US34	17	46	N/A	Nuclear	20 kDa band detected by western blot in lysates of insect cells infected with recombinant baculovirus (Gao et al. 2005).
US34A	8	15	N/A	Sub-cytoplasmic	N/A
TRS1	84	80	Larger and smaller species also detected	Sub-cytoplasmic	Molecular mass of pTRS1 detected by Western blot in lysates of AD169-infected fibroblasts was consistent with the predicted size of 84 kDa (Romanowski and Shenk 1997). pTRS1 was detected in the nucleus and cytoplasm of infected cells at IE times, but predominantly cytoplasmic at late times of infection. Within transfected HeLa cells, pTRS1 was also predominantly cytoplasmic.

3.4.1. Cloning of codon-optimised UL55 and UL74 genes

gpUL55 (gB) and gpUL74 (gO) are well-characterised virion glycoproteins, yet their expression was not detectable when the constructs generated for the library were tested. To address this issue, *de novo*-synthesised, codon-optimised UL55 and UL74 genes were cloned into AdZ by recombineering by Dr James Davies. Following codon optimisation, expression of these genes was readily detected (Figure 3.3). This approach is therefore currently being extended to the additional constructs whose expression was not detectable.

3.4.2. Cytoxic gene products

Infection of HFFF-hCARs with RAds was routinely performed at an MOI=10 to ensure efficient gene delivery and expression. However, the Ad vectors can drive high level gene expression, and HCMV genes can be expected to impact on normal cellular processes. RAd vectors containing the following genes were associated with overt cytopathology: RL8A, RL10, RL11, UL2, UL5, UL6, UL15A, UL16, UL18, UL19, UL20, UL23, UL26, UL34, UL41A, UL138, US14, US26, US28 and US30. To foster survival of the cells for long enough to screen gene function, experiments using these RAd constructs were performed at an MOI=1; little cytotoxicity was observed at MOI=1 (data not shown).

3.4.3. Western blot analysis of expressed HCMV gene products

3.4.3.1. Cross-comparison of western blot data with published findings

Expression of 152 HCMV genes was detected in western blot experiments using a V5 monoclonal antibody, or antibodies specific for HCMV proteins (Figure 3.3). Amongst these, the molecular weight of 16 putative HCMV glycoproteins has not been reported: RL12, UL6, UL7, UL8, UL9, UL10, UL11, UL116, UL120, UL121, UL124, UL147, UL148, US1, US29 and US30 (Table 14.1; Davison and Bhella 2007). Major western blot species migrating more slowly than predicted from primary translation products were detected for 11 of these (RL12, UL6, UL7, UL8, UL9, UL10, UL11, UL116, UL124, UL148 and US29) whilst the other 5 migrated at, or faster than their predicted primary translation products.

Deglycosylation analysis with EndoH and PNGaseF enzymes was performed on products of the first three putative glycoprotein-coding genes, RL12, UL6 and UL7, and the presence of carbohydrate moieties confirmed (Figure 3.4). gpRL12 showed sensitivity to EndoH treatment, reducing in size from 80 kDa to ~50 kDa, close to the 47 kDa predicted size of its polypeptide backbone. Incubation with PNGaseF did not further reduce the molecular weight of gpRL12. Native gpUL6 migrated at 50 kDa, although a weakly expressed 60 kDa band was also seen. The main 50 kDa species was reduced to 28 kDa upon treatment with EndoH, and to 27 kDa after PNGaseF digestion. That both species migrated faster than the predicted 31 kDa primary protein indicates they may undergo some degree of cleavage. The 60 kDa band was completely resistant to EndoH, and partially sensitive to PNGaseF (37 kDa). The primary translation product of UL7 has a predicted mass of 24 kDa. Major and minor bands were observed at 53 kDa and 80 kDa respectively. The 53 kDa isoform was sensitive to EndoH treatment, migrating at 25 kDa, and was further reduced to 24 kDa following digestion with PNGaseF. By contrast, the 80 kDa isoform was unaffected by EndoH, but cleaved to 35-40 kDa by PNGaseF. The observation that predominant western blot species of gpRL12, gpUL6 and gpUL7 are all sensitive to EndoH digestion is consistent with the staining pattern visualised for these products by immunofluorescence. All three proteins accumulated in juxtanuclear sites resembling the ER (see Figure 3.3). This is most clearly seen with gpUL6, which formed a tight ring around the nuclei of infected cells. The presence of weakly-expressed, high-molecular weight gpUL6 and gpUL7 is also interesting. Sensitivity to EndoH and partial sensitivity to PNGaseF indicates passage of these isoforms through the Golgi, and O-linked glycosylation respectively. These heavily glycosylated species may therefore encode distinct functions compared with the predominantly expressed variants.

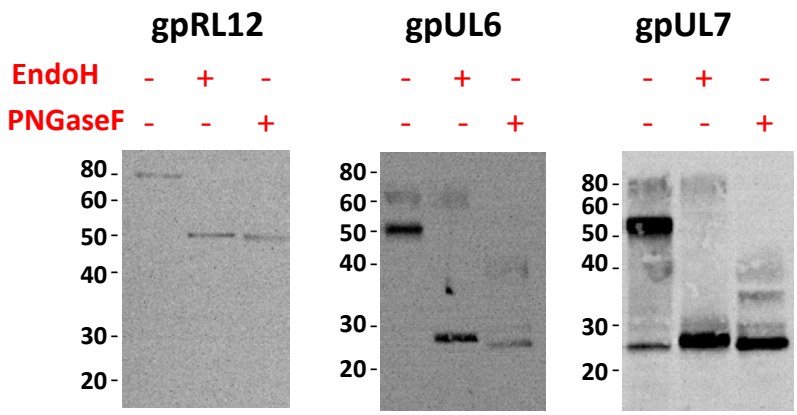


Figure 3.4: Deglycosylation analysis of gpRL12, gpUL6 and gpUL7
Western blot was performed as described in figure 4.3, on RAd-expressed gpRL12, gpUL6 and gpUL7 in native form, or following treatment with EndoH or PNGaseF.

The remaining 136 genes include 116 for which product sizes have been reported (see Table 3.2), and 20 for which predicted sizes were calculated from amino acid sequence. Major western blot species were detected within 10 kDa of the expected size for products of 128 genes. Of the other 8 (UL28/29, UL56, UL82, UL87, IRS1, US18, US19 and US34; indicated in bold in table 3.2), the first 5 encode products in excess of 70 kDa, and had calculated mass differences of less than 15 kDa. These differences may be attributable to the accuracy with which SDS-PAGE can resolve heavier proteins at the top of the gel. Products with approximately appropriate sizes were observed for US18 and US19, however in the case of US18 these were accompanied with more abundant high molecular weight species (see section 3.4.3.5). Finally, US34 migrated at 46 kDa, almost 30 kDa higher than its predicted molecular weight of 17 kDa. Although originally not predicted to be a glycoprotein by Davison and Bhella (2007), the US34 ORF contains 4 potential N-linked glycosylation sites, all of which are conserved amongst HCMV strains, and 3 of which which are conserved in CCMV (Dr Andrew Davison, personal communication). High-scoring potential phosphorylation sites are also present in the US34 ORF. US34 may therefore encode a glycoprotein, or a phosphorylated protein.

3.4.3.2. Altered processing of virion glycoproteins during productive HCMV infection

A subset of HCMV gene products can be expected to exhibit a different size to that calculated from its primary translation product when expressed individually from a vector. This is because an HCMV protein expressed individually loses the capacity to interact with any other viral counterpart(s) that may induce post-translational modification. Some of these modifications, particularly glycosylation and/or proteolytic cleavage, clearly have the capacity to dramatically alter the size of the protein detected. Western blot analysis of RAd-derived HCMV genes recapitulated reports of altered molecular weight (Table 3.2); some examples are included below.

In HCMV virions, production of functional gCI complexes requires the cleavage of 160 kDa gB precursors into 116 kDa-55 kDa mature complexes. Whereas in HCMV-infected cells the 55 kDa isoform is most abundantly detected, expression in isolatation results in the accumulation of the uncleaved precursor. My observation that the predominant RAd-expressed gB isoform migrated at 150 kDa on a reducing gel is thus consistent with published findings. While HCMV gCII consists of a gM-gN complex, gN expressed alone was reported to have a relative molecular mass of 18 kDa (Mach et al., 2000), close to the predicted size of 14 kDa. However when co-expressed with gM, glycosylation of gN increases, resulting in migration at 50-60 kDa.

Consistent with these findings, gN expressed using AdZ migrated most abundantly at 17 kDa, although the detection of additional slower migrating species of up to 34 kDa indicated limited gM-independent maturation of gN. Classically, gCIII was recognised to consist of the gH/gL/gO trimeric complex. As observed with gN, gO was reported not to be fully processed to its mature form (100-125 kDa) when expressed independently (Theiler and Compton, 2002). Detection of a major 85 kDa species from our RAd-UL74 construct is in agreement with this finding.

As shown in Figure 3.5, 55 kDa and 80 kDa protein species were expressed from RAd-RL13, whilst 55 kDa and 100 kDa gpRL13 products were detected during productive HCMV infection. The 55 kDa species expressed from both viruses were fully sensitive to EndoH digestion, indicating ER-retention. The 80 kDa and 100 kDa band proteins were resistant to EndoH treatment, and only partially sensitive to digestion by PNGaseF, suggesting mature RL13 contains O-linked, as well as N-linked carbohydrate moieties.

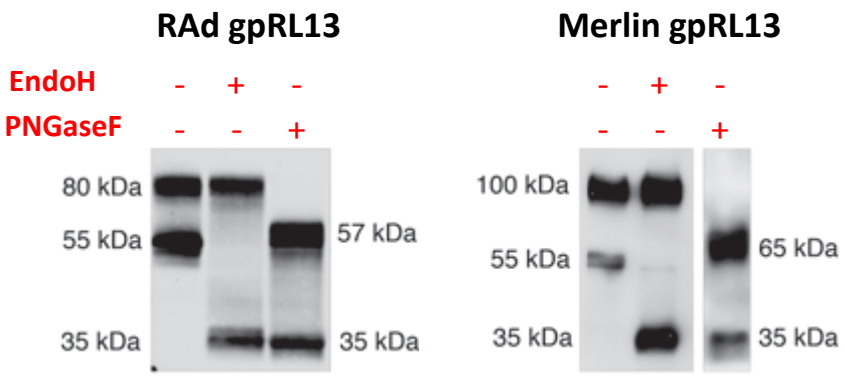


Figure 3.5 **(adapted from stanton et al. 2010):**
Glycosylation analysis of gpRL13 expressed from a RAd, and in the context of HCMV-Merlin infection
Western blot performed on RAd-expressed, and HCMV Merlin-expressed gpRL13 in native form, and following digestion with EndoH or PNGaseF

3.4.3.3. Migration of HCMV-encoded phosphoproteins in western blot

14 HCMV genes are known to encode phosphoproteins (green cells in Table 3.2; UL32 encodes a phosphoglycoprotein). The majority of HCMV-encoded phosphoproteins are components of the virion tegument (Kalejta, 2008). The fact that most tegument proteins exhibit some degree of phosphorylation has lead to the proposal that this modification could promote their recruitment to this virion compartment; no specific sequence has been identified that targets tegument constituents to sites of virus assembly. In contrast with glycosylation, phosphorylation rarely results in significant increase in molecular weight. Consistent with this notion, RAd-encoded phosphoproteins generally migrated at near their predicted/published molecular mass (green cells, table 3.2). Interaction with other HCMV proteins is therefore unlikely to result in significant alteration to the molecular weight of phosphoproteins.

UL112 encodes four phosphoproteins with common N-termini, which are synthesised from a single alternatively spliced transcript (Wright and Spector, 1989). These phosphoproteins migrate at 34 kDa, 43 kDa, 50 kDa and 84 kDa in the context of HCMV infection, and almost identically-sized products were expressed from RAd-UL112. Also consistent with previous reports, the 43 kDa isoform was the most abundant of the four species. This demonstrates that neither alternative splicing, nor phosphorylation of UL112 gene products requires the presence of other viral proteins.

3.4.3.4. US12 family of HCMV proteins

The US12 family consists of 10 adjacent genes positioned between US12 and US21 inclusively, all of which remain incompletely characterised. These genes have been predicted to encode seven-transmembrane-domain proteins (McGeoch et al., 1985), and also show features reminiscent of G-protein-coupled receptors (Lesniewski et al., 2006). Expression from RAd-US14 and RAd-US17 was not detected, whilst pUS19 migrated at 21 kDa, somewhat faster than the 32 kDa protein reported during HCMV infection (Guo and Huang, 1993). The other 7 family members (US12, US13, US15, US16, US18, US20, US21) all provided multiple positive signals ranging from their predicted protein sizes to very much larger sizes, up to >260 kDa. Interestingly, size valuation of some of these proteins in the context of HCMV infection (US16, US18, US19, US20) tended to be more in line with their predicted sizes (see Table 3.2). This difference implies that this set of proteins is subject to a high level of post-translational

modification when expressed in isolation, which is suppressed in the context of productive virus infection. The high-molecular weight species of the aforementioned US12 family members bear resemblance to the migration pattern observed for ubiquitin-conjugated proteins. In support of this notion, immunoprecipitation of pUS20 from lysates of hFFF-hCARs infected with RAd-US20 using the V5 antibody, followed by stable isotope labelling by amino acids in cell culture (SILAC) analysis revealed pUS20 interacts with ubiquitin, p62 and Tax1-binding protein 1 (TAX1BP1) (Dr Ceri Fielding, personal communication). p62 is involved in the targeting and degradation of polyubiquitinated proteins by proteasome and autophagy (Seibenhener et al., 2004, Myeku and Figueiredo-Pereira, 2011) whilst TAX1BP1 is important in ubiquitin-mediated signalling, and may also act as a ubiquitin-binding protein adaptor (Verstrepen et al., 2011). A subseqeunt study by Dr Ceri Fielding has since demonstrated experimentally that pUS20 is indeed polyubiquitinated (personal communication). It is predicted that in addition to pUS20, the other 6 US12 family members identified above are also ubiquitinated during ectopic expression. This is currently being investigated. Interestingly, the HCMV matrix protein UL48 has been shown to be a deubiquitinase (Adair et al., 2002), and thus in lytic infection potentially could act to limit post-translational modification of US12 family proteins.

3.4.3.5. Efficiency of expression

It is noteworthy that expression level varied dramatically between different Ad constructs; appropriate exposure of western blots was selected so as not to under or over-expose bands. The presented results are therefore not quantitative, and should not be interpreted that way. Most notably, lysates of cells infected with a subset of RAds were diluted prior to electrophoresis due to very strong expression. Compared to the standard protocol specified in section 2.11.2, lysates of cells infected with RAds expressing UL14, UL21A, UL26, UL31 and UL38 were diluted 100 times, whilst those of UL123, UL138 and UL150 were diluted by a factor of 10. 1 in 5 dilutions were prepared for RAds encoding UL25 and UL28/29; and RAd-UL112 and RAd-UL114 samples were diluted 20 times and 50 times respectively. By contrast, expression of the following 14 genes was relatively weak in western blot experiments: UL18, UL19, UL30, UL41A, UL48A, UL51, UL72, UL73, UL80, UL80.5, UL117, UL128, UL142 and UL148D. To visualise products of these genes, protein concentration or antibody dilution were not adjusted; blots were instead exposed for prolonged lengths of time where necessary.

3.4.3.6. Western blot analysis of RAd-infected-cell supernatants

Western blot was also performed on the supernatants of RAd-infected cells in order to identify novel secreted gene products. Abundant levels of gpUL4 were detected in the culture medium of cells infected with RAd-UL4 using the V5 antibody (see Chapter 6).

3.4.4. Intracellular localisation of RAd-expressed HCMV gene products

3.4.4.1. Cross comparison of immunofluorescence data with published findings

Expression of 155 HCMV genes was detected by immunofluorescence using a V5 monoclonal antibody, or antibodies specific for HCMV proteins (Figure 3.3). Overall, intracellular localisations correlated well with information in the published literature (Table 3.2). Localisations have been reported for 109 of the 155 detected gene products. Data presented here closely resembled previous reports in the case of 98 of these, and varied in the case of 11 others (underlined in table 3.2). Expression was categorised as inconsistent with the literature if it was (i) exclusively nuclear but had been reported as cytoplasmic or pan-cellular, (ii) exclusively cytoplasmic but had been reported as nuclear or pan-cellular, (iii) pan-cellular but had been reported to be either nuclear or cytoplasmic, (iv) pan-cytoplasmic but had been reported as sub-cytoplasmic, (v) sub-cytoplasmic, but had been reported as pan-cytoplasmic/cell surface, (vi) nuclear but had been reported as sub-nuclear, (vii) sub-nuclear but had been reported throughout the nucleus. Where distinct sub-cellular localisations had been reported during HCMV infection and isolated expression (see Section 3.4.4.4), results were compared to the latter group, except if isolated expression was visualised in non-human cells.

Immunofluorescence staining was inconsistent with previous literature for products of the following 11 genes: UL38, UL43, UL51, UL77, UL80.5, UL85, UL89, UL98, UL104, UL130 and US22. Compared to previous descriptions, pUL43, pUL80.5 and pUL130 were found in the correct compartment (nucleus/cytoplasm), but distributed differently within that compartment. With RAd-UL43, expression was visualised throughout the cytoplasm including the cell surface. pUL43 is a tegument protein previously reported to localise to juxtanuclear virus assembly compartments during lytic infection (Adair et al., 2002). When expressed in isolation virus assembly compartments are not generated, and pUL43 becomes diffusely

cytoplasmic. RAd-expressed pUL80.5 displayed punctate nuclear staining, but has been reported to be expressed throughout the nucleus of transfected Cos-7 cells (Plafker and Gibson, 1998, Wood et al., 1997). Isolated expression of pUL80.5 in human cells has not been reported; the difference in pUL80.5 nuclear localisation could therefore be species-specific. RAd-UL130 expression was detected at sub-cytoplasmic regions of infected cells. Gerna et al. reported cell surface, as well as cytoplasmic localisation for pUL130 transiently expressed in 293T cells (Gerna et al., 2008b). It is possible that pUL130 localisation varies depending on the cell type used.

Interestingly, localisations of products of all other 8 genes (UL38, UL51, UL77, UL85, UL89, UL98, UL104 and US22) were (i) previously reported to be nuclear or predominantly nuclear, but observed to be cytoplasmic in Figure 3.3, and (ii) previously described solely in the context of HCMV infection. It is therefore likely that they depend on interaction with other viral proteins for nuclear targeting. In this regard, 5 of the aforementioned 8 gene products are found in purified virions (UL38, UL77, UL85, UL98 and US22 (Lahijani et al., 1992, Varnum et al., 2004), whilst pUL104 is the portal protein of pro-capsids (Dittmer and Bogner, 2005). Components of the virion often form complexes which may be important in determination of localisation.

3.4.4.2. Gene products that require other HCMV proteins for appropriate localisation

A number of HCMV gene products have been reported to traffic inappropriately within the cell when expressed individually. The specific virus-encoded interacting partners required for appropriate translocation of some of these gene products are known, whilst others await identification. Immunofluorescent detection of HCMV genes by using AdZ recapitulated published findings for this group of proteins, although rescue of localisation by co-expression with interacting partner genes was not performed.

HCMV gene products that localise differently in the absence of known viral co-factors include pUL86 (MCP) which depends on the presence of pUL80.5 (pAP) for nuclear translocation, pUL53 which requires co-expression with pUL50 for trafficking out of the nucleus and into the nuclear membrane, and pUL94 which is exclusively nuclear in the absence of pp28, but pan-cellular when co-expressed with pp28. In addition, gH and gO both need co-expression with gL for movement to the cell membrane whilst gM and gN are localised in the ER and ERGIC during ectopic expression, but extended to the cellular secretory pathways when co-expressed.

HCMV proteins that localise to inappropriate cellular compartments during ectopic expression can also be expected to display aberrant function when expressed alone. This has potentially important implications for use of the RAd library in functional screens, and is discussed further in chapter 7.

3.4.4.3. Dependence of intracellular localisation on MOI/expression level

To optimise conditions for image collection, cells were expressed at a range of MOIs, and it was clear that the level of expression correlated with the amount of virus in the inoculum. Moreover, the intracellular localisation of some proteins was dependent on the MOI used for RAd-infection; more specifically, pUL26, pUL27 and pUL28/29 were detected exclusively in the nucleus when using an MOI of 1 (Figure 3.6). However, an increase to 5 or 10 pfu per cell was associated with cytosolic expression of all three proteins and induction of cytotoxicity. At an MOI of 5, UL26 expression was less conspicuous within nuclei, and more pronounced in the cytoplasm. Over-expression of pUL27 (10 pfu per cell) led to apparent chromatin condensation, as indicated by brighter and more focused DAPI staining. In these cells pUL27 localised in ring-like structures around the nuclei. pUL28/29 expression at MOI 10 caused apparent loss of integrity to the nuclear membrane. UL26, UL27 and UL28/29 encode proteins that adopt nuclear localisations in the context of HCMV infection. UL26 encodes a tegument protein that can transactivate the IE-promoter (Stamminger et al., 2002), whilst pUL27 is responsible for Maribavir resistance, and may play a role in viral replication and egress (Komazin et al., 2003). Nuclear and nucleolar expression has previously been reported for a UL27-GFP fusion protein in transfected fibroblasts (Chou et al., 2004). pUL28/29 is the product of a spliced transcript encoded by UL28 and UL29, present within viral replication compartments in infected cell nuclei, where it may play a role in regulating viral gene expression (Mitchell et al., 2009).

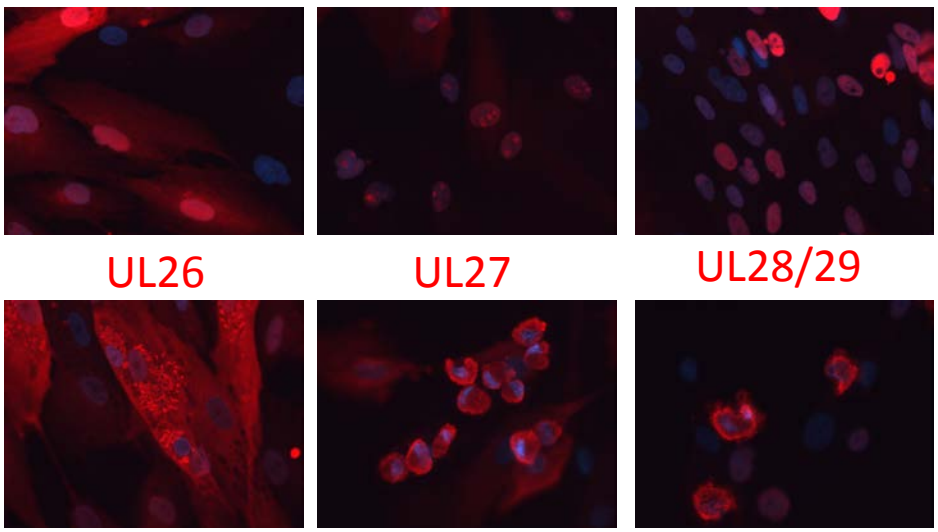


Figure 3.6: Effect of transgene expression level on protein localisation
HFFF-hCARs were infected with RAd-UL26, RAd-UL27 and RAd-UL28/29 at MOI 1 (top panel) or MOIs 5, 10 and 5 (bottom panel) respectively. Immunofluorescence staining was performed at 3 days p.i. As described in figure 4.3.

Although the AdZ system is recognised to be capable of providing for very high level expression, there was substantial variation among the 170 different constructs analysed. Over-expression has the potential to saturate cellular pathways, and thereby result in inappropriate protein trafficking and toxicity. The clearest and most specific demonstration of intracellular localisation tended to be observed when expression levels were modest, and that is most clearly demonstrated with UL26, 27 and UL27/UL28. For immunofluorescence data collated in Figure 3.3 and table 3.2, infections were predominantly performed at the lowest MOI compatible with detecting expression. These results also provide a message for high throughput functional screening in that over expression of certain genes could perturb their function.

3.4.5. Constructs lacking evidence of expression

Since the ß2.7 gene encodes an abundant transcript without an obvious ORF, the RAd-β2.7 construct was not compatible with epitope-tagging and not shown in Figure 3.3. Expression from 8 tagged constructs encoding RL6, RL9A, UL30A, UL33, UL146, US9, US14 and US17 was not detected by either western blot or immunofluorescence. It is possible however, that these proteins expressed a product that was not detected using the V5-specific antibody. Detection of protein expression requires that the C-terminal V5 tag be accessible to the antibody, and can be prevented by ill-suited protein folding. Alternatively, C-terminal cleavage of the recombinant protein would impede detection of expression, though this is unlikely to be the case for protein products of RL6, UL33, US9 and US14, since these proteins have previously been detected using C-terminal epitope tags (Salsman et al., 2008). RL9A was recently annotated as a protein-coding gene in the Merlin genome, being added together with RL8A, UL150A and US33A (Gatherer et al., 2011). RAd-RL8A expression was detcted by immunoblotting but not immunofluorescence, whilst UL150A was not cloned into AdZ due to uncertainties about the 5' start site of its major splice variant (see chapter 5). UL30A is related to the neighbouring UL30 gene, and is conserved in primate CMVs (Davison, 2010). UL30A is predicted to use ACG instead of the conventional ATG as a translation initiation codon; it is possible that the non-canonical start codon was not functional in the AdZ vector. It is also possible that UL30A could be expressed if supplied with a conventional start codon; this is being addressed. Expression of genes RL5A, UL1, UL76, UL77, UL100, UL132, UL139 and UL148A was not detected by western blot, and only weakly detected by immunofluorescence; the threshold of expression for detection using the V5 antibody may therefore be lower by immunofluorescence compared to immunoblotting. The products of RL8A, UL142, UL147A,

US1, and US33A were only detected by western blot. One explanation for this may be that antibody binding occurred under denaturing and reducing conditions, but not when the protein was in native form.

3.5. Summary

168 protein-coding genes encoded by HCMV strain Merlin were cloned into Ad vectors to generate a high-throughput system for use in functional assays (see Table 3.1 for individual contributions). Expression of 160 genes was verified by western blot (152 genes) and immunofluorescence (155 genes) whilst codon-optimised versions of the other 8 genes are currently being cloned into AdZ to facilitate their expression. This is the first systematic analysis of wildtype HCMV gene expression reported to date. Expression data obtained for the great majority of HCMV gene products was in accord with previous reports, indicating that this library of Ad recombinants can be used with confidence in functional assays.

Expression of ORFs encoded by HCMV strains AD169 and Towne (ULb' ORFs) was undertaken by using a baculovirus vector (Gao et al., 2005), and by transient DNA-transfection of 293T cells and osteosarcoma cells (Salsman et al., 2008). The Gao et al. paper did not contain the primary data but provided the size of all expressed gene products as obtained by western blot; 50% of genes tested provided a detectable product. In contrast, transient transfection of the 293T cells tracked expression mainly by immunofluorescence only. In the case of non-glycosylated proteins, excellent overall co-linearity was observed between data provided by Gao et al. and results presented here. However, the baculovirus expression system did not faithfully produce glycoproteins of appropriate size. Glycoproteins consistently migrated at or near the predicted size of their primary translation products; indeed baculovirus vector-driven expression in insect cells is known to cause alterations to glycosylation pathways. It was therefore not beneficial to reference glycoprotein sizes from the Gao et al. report in table 3.2. There was broad agreement linking immunofluorescence data obtained using Ad recombinants and the Salsman et al. (2008) report, although one major inconsistency was observed between the two datasets. Several gene products that show specific sub-cellular localisation when expressed from Ad recombinants were classified as pan-cellular by Salsman and co-workers. These included products that are known not to be expressed throughout the cell: UL20, UL24, UL43, UL56, IE2, UL94, UL114, UL138, US24 and US28 gene products (see Table 3.2). Relatively few HCMV genes have been detected as showing pan-cellular expression by others and myself. It is likely that pan-cellular localisation of some proteins observed by Salsman et al. was due to

over-expression of HCMV proteins; this can lead to lack of containment within appropriate cell compartment(s).

Expression analysis of the RAd library of HCMV genes led to characterisation of the glycosylation state of 3 predicted glycoproteins (RL12, UL6 and UL7), identification of gpUL4 as a potential secreted protein, and identification of pUS12, pUS13, pUS15, pUS16, pUS18, pUS20, and pUS21 as potential ubiquitin-binding proteins.

Expression with 20 Ad recombinants was associated with visible cytopathic effect at MOI=10. To avoid this bottleneck, it was decided to administer these Ad recombinants at MOI=1 in functional assays. In addition, the localisation of some proteins depends on the level of expression, as infection with their containing RAds at a range of MOIs resulted in significant variation in transgene distribution within the cell. This was most conspicuous in the cases of pUL26, pUL27 and pUL28/29; these proteins were all nuclear at MOI=1, consistent with their reported expression during productive HCMV infection. Increase of MOI to 5-10 however resulted in segregation from nuclei, as evidenced by lack of co-localisation with DAPI. These examples highlight the need to consult independent reports that describe localisation, prior to prescribing MOIs of Ad recombinants for use in functional assays.

Ectopic expression of some HCMV gene products results in altered localisation and/or molecular mass. Such alterations may also be accompanied by abrogated function, because appropriate trafficking and post-translational modifications are vital to protein function. Specific HCMV proteins required for appropriate translocation and/or glycosylation of some gene products are known; co-expression of these ‘mediator’ proteins together with their target proteins can be beneficial during functional assays. Infection of cells with more than one Ad recombinant in this context may however increase cytotoxic affects associated with each Ad recombinant.

Where the viral function of interest necessitates the concerted action of two or more genes, gain of function screening of single genes is rendered ineffective. In these instances, co-infection of cells with multiple Ad recombinants is impractical due the extremely high number of possible combinations (>28,000 pair-wise combinations for 168 Ad recombinants).

The majority of genes were expressed with C-terminal tags. Obvious benefits of this approach include the potential to rapidly and efficiently (i) assess gene expression, (ii) purify tagged proteins, and (iii) study protein-protein interactions by immunoprecipitation. A disadvantage of this epitope-tagging strategy lies in uncertainties about interference of the tag with protein function. For a small number of genes such as the MHC class-I-downregulating US2, US3, US6

and US11 genes, function of the tagged protein can be tested with ease (see chapter 4), but this is not possible for most other gene products.

Use of the RAd library in functional screening of HCMV genes is detailed in chapter 4, and discussed further in chapter 7.

4- Regulation of Fas expression by HCMV

4.1. HCMV and death receptors

HCMV encodes functions that act to suppress both intracellular, and death receptor-induced apoptotic signalling (see Section 1.7). The effect of HCMV infection on the surface expression levels of the death receptor molecules has not been fully elucidated. Although infection with strain AD169 results in reduced TNFRI expression at the cell surface, the HCMV gene(s) responsible has yet to be identified (Baillie et al., 2003). Interestingly, strains carrying an intact ULb' region do not downregulate cell surface TNFRI (Montag et al., 2006). This is because a ULb' gene (UL138) potentiates cell surface TNFRI expression, thus masking the TNFRI-downregulating function (Le et al., 2011, Montag et al., 2011).

Fas (CD95, Apo-1) is a member of the TNF receptor superfamily that is capable of promoting apoptosis following stimulation by its paired activator, Fas ligand (Kägi et al., 1994). The HCMV UL36 and UL37 gene products efficiently inhibit Fas-mediated apoptosis by inhibiting caspase-8 activation and cytochrome C release respectively (Goldmacher, 2005, Skaletskaya et al., 2001). Nevertheless the effect of HCMV infection on expression of Fas itself is poorly understood. Chaudhuri et al reported that infection with the Smith and Eisenhardt strains stimulated Fas surface expression in fibroblasts leading to their killing, and that Fas could not be detected on uninfected fibroblasts (Chaudhuri et al., 1999). This is somewhat surprising given that Fas expression has routinely been observed on human fibroblasts by others (Aggarwal et al., 1995, Freiberg et al., 1997, Jelaska and Korn, 1998, Tepper et al., 2000). Also, considering HCMV elicits broad anti-apoptotic responses, upregulation of Fas in infected cells is counter-intuitive. I therefore sought to investigate the regulation of Fas surface expression by HCMV.

4.2. HCMV downregulates Fas from the cell surface

Fas expression on the surface of strain Merlin-infected and mock-infected fibroblasts was analysed by flow cytometry in a time-course assay. As shown in Figure 4.1, Fas was consistently downregulated on infected cells at 1, 2, and 3 days p.i.. MHC class-I staining was included as a positive control for the efficiency of infection. Genes US2, US3, US6 and US11 are all known to decrease MHC class-molecules at the surface of infected cells (see Section 1.7.3.2). Consistent with this expectation, MHC class-I was efficiently downregulated at all three time-points. Comparative levels of Fas downregulation were also observed at the true-late stage of infection (6 days p.i., data not shown). Downregulation of cell surface Fas is a

novel HCMV function, with potential implications in the interaction between HCMV infected cells and the host's immune system.

4.2.1. Screening a panel of Merlin block deletion mutants for loss of Fas downregulation

I next aimed to map the gene responsible for Fas downregulation by screening a panel of Merlin block deletion mutants in a loss of function assay. Figure 4.2 is a diagrammatic representation of the region deleted in each of these mutants. The block deletion mutants were all produced in a Merlin ΔUL16, ΔUL18 virus carrying a GFP tag on the end of UL32 (designated RCMV1278). RCMV1278 was initially designed as a backbone for block deletions because of its utility in NK cell modulation. *In vitro*, cells infected with wild-type Merlin are almost completely resistant to NK cell recognition, whereas uninfected cells stimulate NK cell activation. UL16 and UL18 are known to contribute to NK cell suppression; infection with RCMV1278 therefore results in intermediate NK cell inhibition. Deletion of further NK cell evasion genes from 1278 would then result in a virus more sensitive to NK-mediated killing, whereas deletion of NK cell activating genes would render the virus more resistant to NK cells.

It was necessary to first compare Fas cell surface expression between cells infected with Merlin and RCMV1278. Comparable levels of Fas and MHC class-I downregulation were observed with both viruses at 3 days p.i. (Figure 4.3a). This result established that UL16 and UL18 were not required to suppress Fas expression. GFP was observed by immunofluorescence in cells infected with RCMV1278, but not in mock-infected, or Merlin-infected cells (Figure 4.3b).

The block deletion mutants were then tested alongside 1278 by flow cytometry at 3 days p.i.. Please note that since deletion of some genes elicits growth defects, the MOIs used in this experiment were adjusted, in order to achieve equivalent cytopathic effect, UL32-GFP expression and MHC class-I down-regulation. Wildtype-like downregulation of cell surface Fas was observed with all deletion mutants tested (Figure 4.4). Also, MHC class-I was downregulated by all viruses except the ΔUS2-US11 mutant, which upregulated MHC class I. This upregulation was expected, and is attributed to enhanced interferon signalling in HCMV-infected cells. In this loss-of-function assay, none of the 58 genes deleted from this bank of HCMV mutants was found to be required for Fas downregulation.

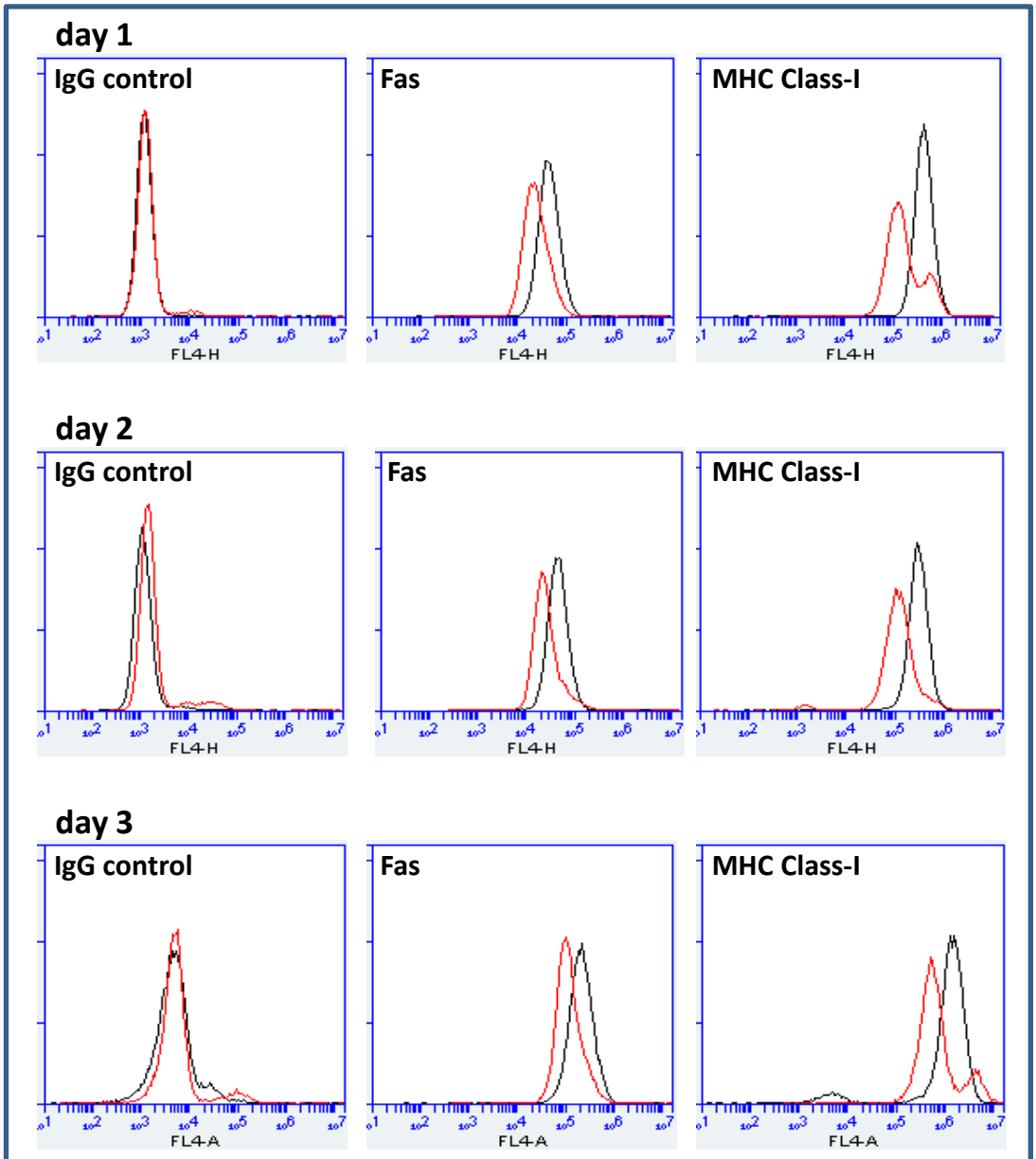


Figure 4.1

Time-course of Fas downregulation by Merlin

Fas expression on the surface of mock-infected and Merlin-infected cells (MOI 10) was analysed by flow cytometry at 1, 2, and 3 days p.i. Cells were stained with Fas, MHC class-I or IgG control primary antibodies, and an Alexa fluor 647 secondary antibody. For each time-point, control IgG, Fas and MHC Class-I staining are shown in the left-hand, middle and right-hand histograms respectively. Mock and Merlin samples are shown in black and red respectively.

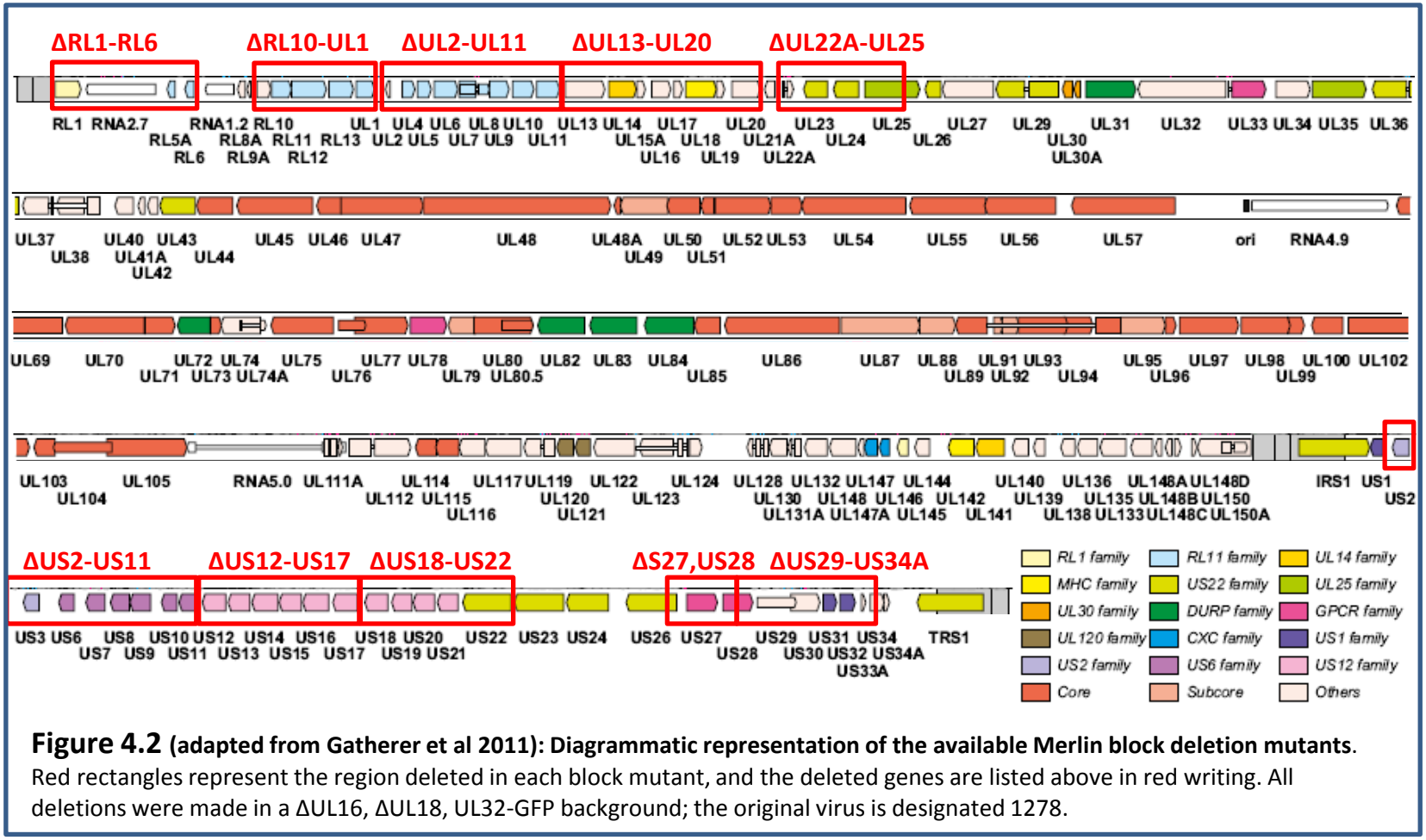


Figure 4.2 (adapted from Gatherer et al 2011): Diagrammatic representation of the available Merlin block deletion mutants. Red rectangles represent the region deleted in each block mutant, and the deleted genes are listed above in red writing. All deletions were made in a ΔUL16, ΔUL18, UL32-GFP background; the original virus is designated 1278.

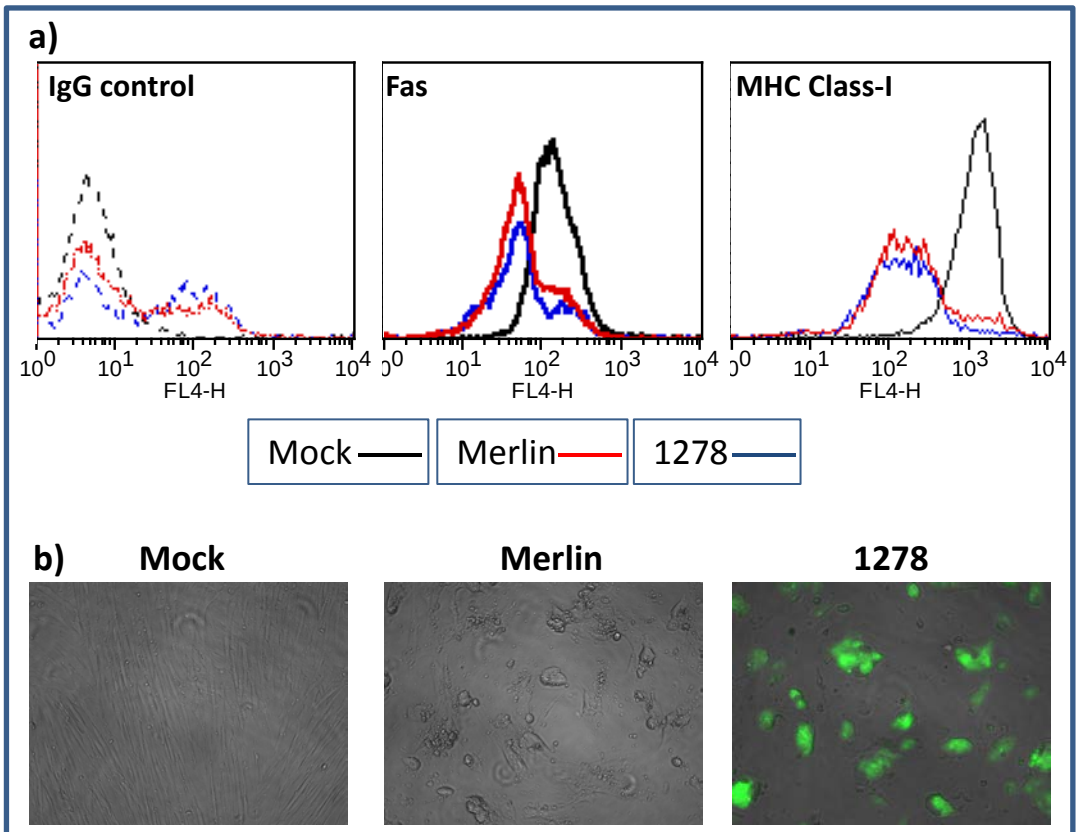
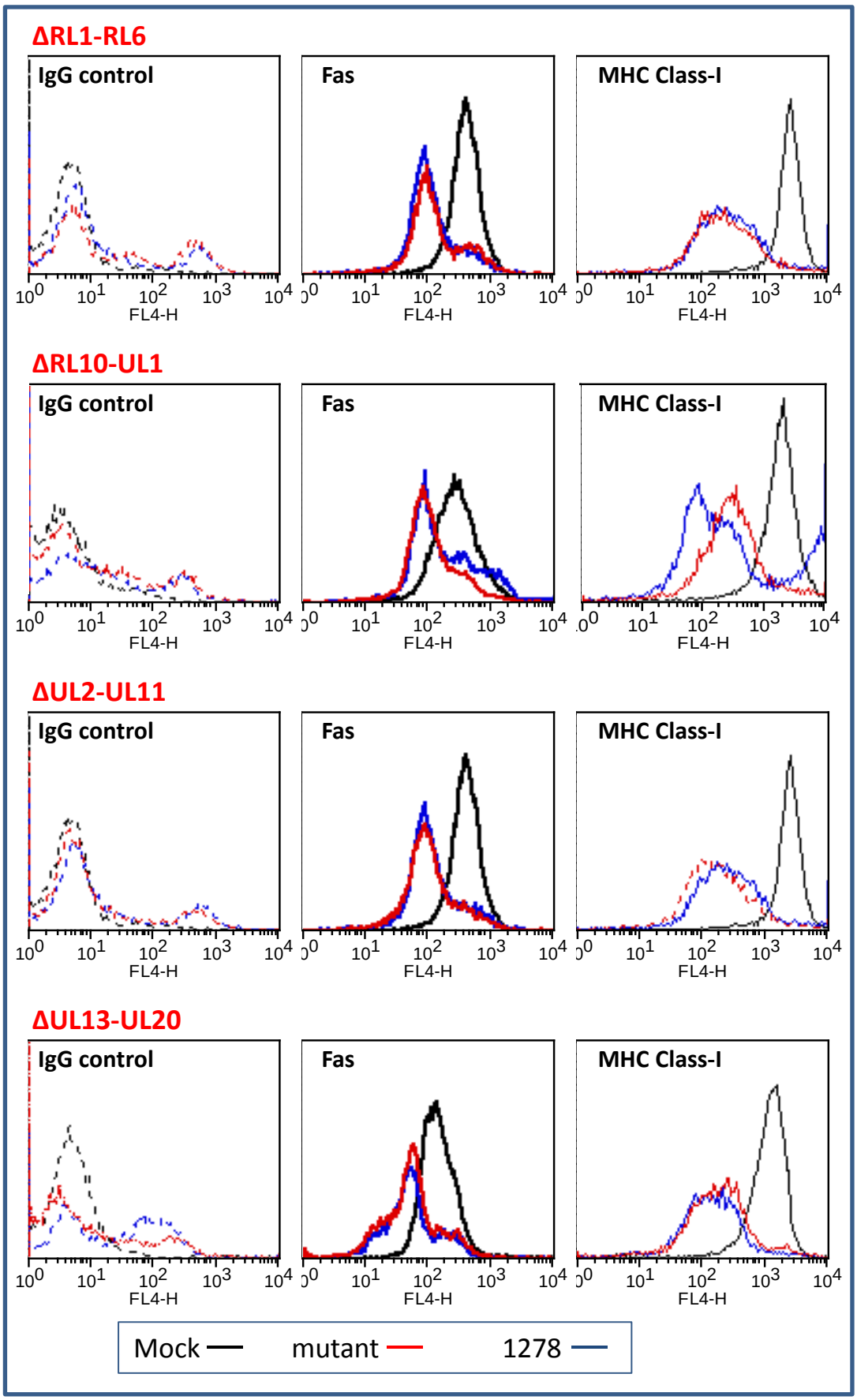
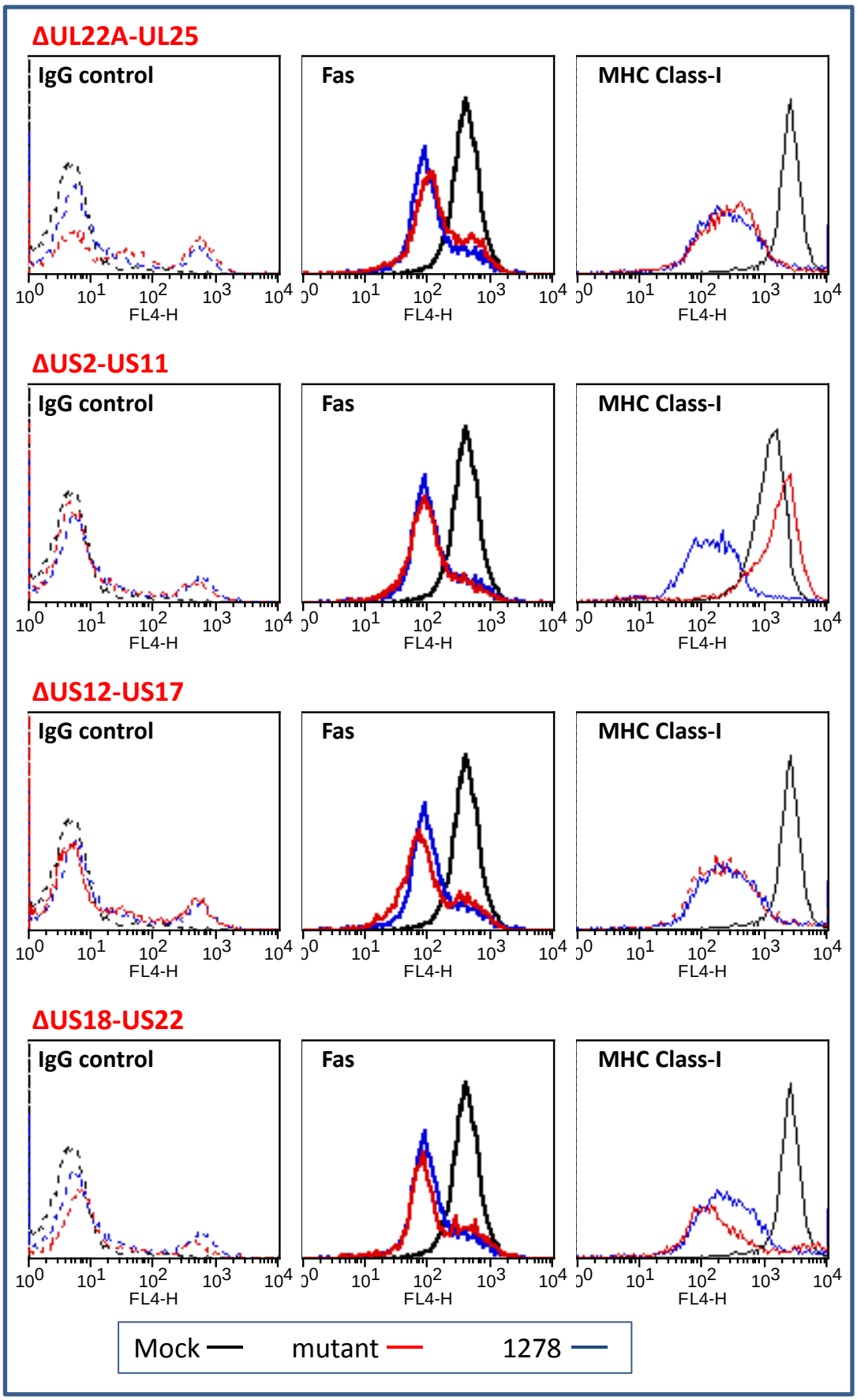


Figure 4.3

Fas downregulation by Merlin and a Merlin ΔUL16, Δ UL18 virus

a) Fas expression on the surface of cells infected with Merlin and a Merlin ΔUL16, ΔUL18, UL32-GFP virus (designated 1278) were analysed by flow cytometry. RCMV 1278 was used to generate Merlin block deletion mutants which were used in loss of function screening. HFFFs were infected at MOI 10 with Merlin and 1278. 3 days post-infection cells were stained with Fas, MHC class-I or IgG control primary antibodies, and an Alexa fluor 647 secondary antibody. Control IgG (dotted lines), Fas (thick lines) and MHC Class-I (thin lines) staining are shown in the left-hand, middle and right-hand histograms respectively. Mock samples are shown in black, whilst Merlin and 1278 samples are displayed in red and blue respectively. b) Images of cell monolayers were obtained prior to harvest of samples; UL32-GFP expression was detected in 1278-infected cells and overlaid with phase contrast image.


ΔRL1-RL6
IgG control
Fas
MHC Class-I
ΔRL10-UL1
IgG control
Fas
MHC Class-I
ΔUL2-UL11
IgG control
Fas
MHC Class-I
ΔUL13-UL20
IgG control
Fas
MHC Class-I
FL4-H
Mock
mutant
1278


ΔUL22A-UL25
IgG control
Fas
MHC Class-I
ΔUS2-US11
IgG control
Fas
MHC Class-I
ΔUS12-US17
IgG control
Fas
MHC Class-I
ΔUS18-US22
IgG control
Fas
MHC Class-I
FL4-H
Mock
mutant
1278

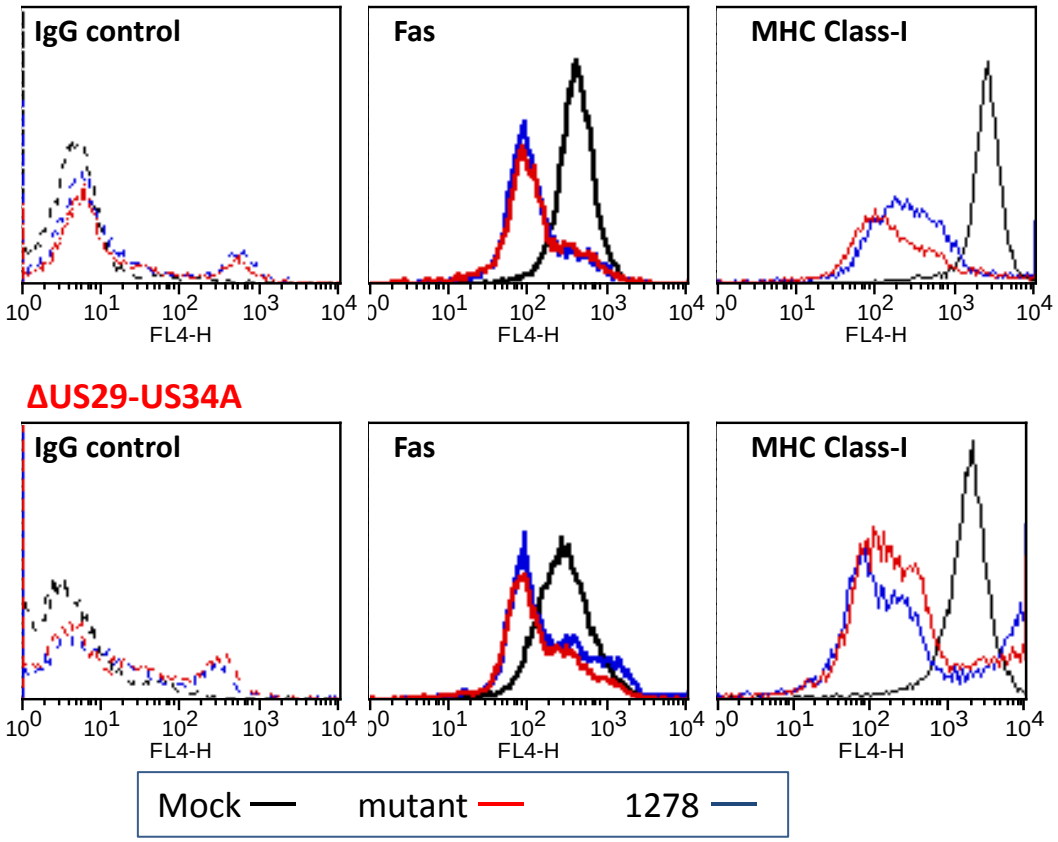


Figure 4.4

Loss of function screening of Merlin block deletion

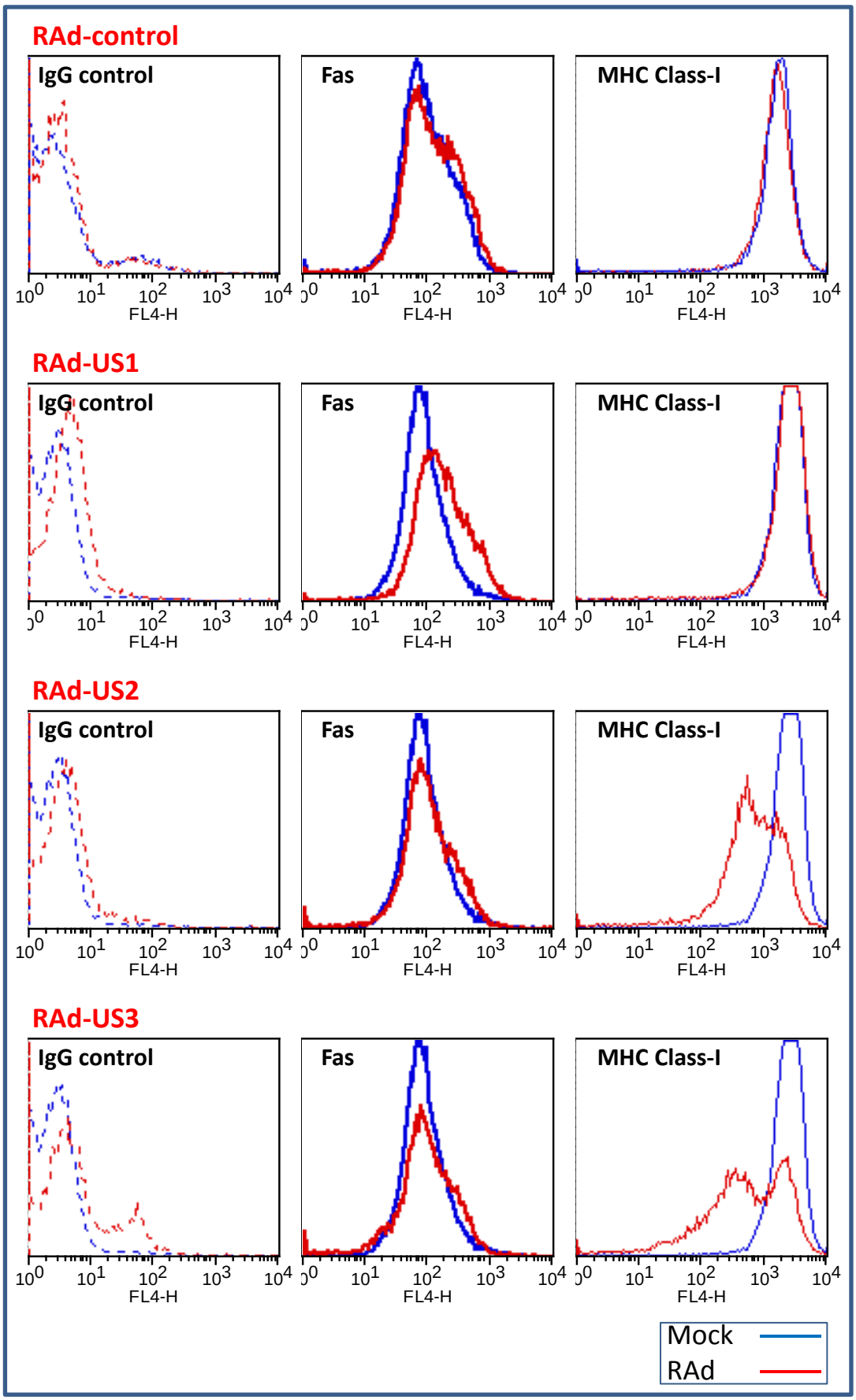
Surface Fas expression was compared by flow cytometry between cells infected with 1278 and various Merlin mutants each lacking a block of non-essential genes. The region deleted in each block deletion mutant is indicated in red writing. Infections were carried out in HFFFs at MOI 10, excepting ΔUL13-UL20 and ΔUS27-US29 viruses were used at MOI 50, and ΔUS29-US34A, ΔRL1-RL6 , and ΔUS18-US22 infections performed at MOIs 15, 25 and 60 respectively. Cells were harvested 3 days p.i. Infected cells were stained with Fas, MHC class-I or IgG control primary antibodies, and an Alexa fluor 647 secondary antibody. Control IgG (dotted lines), Fas (thick lines) and MHC Class-I (thin lines) staining are shown in the left-hand, middle and right-hand histograms respectively. Black lines represent mock samples, whilst 1278 and deletion mutant samples are displayed in blue and red respectively. Cells infected with 1278 or any of the deletion mutants were gated for GFP expression.

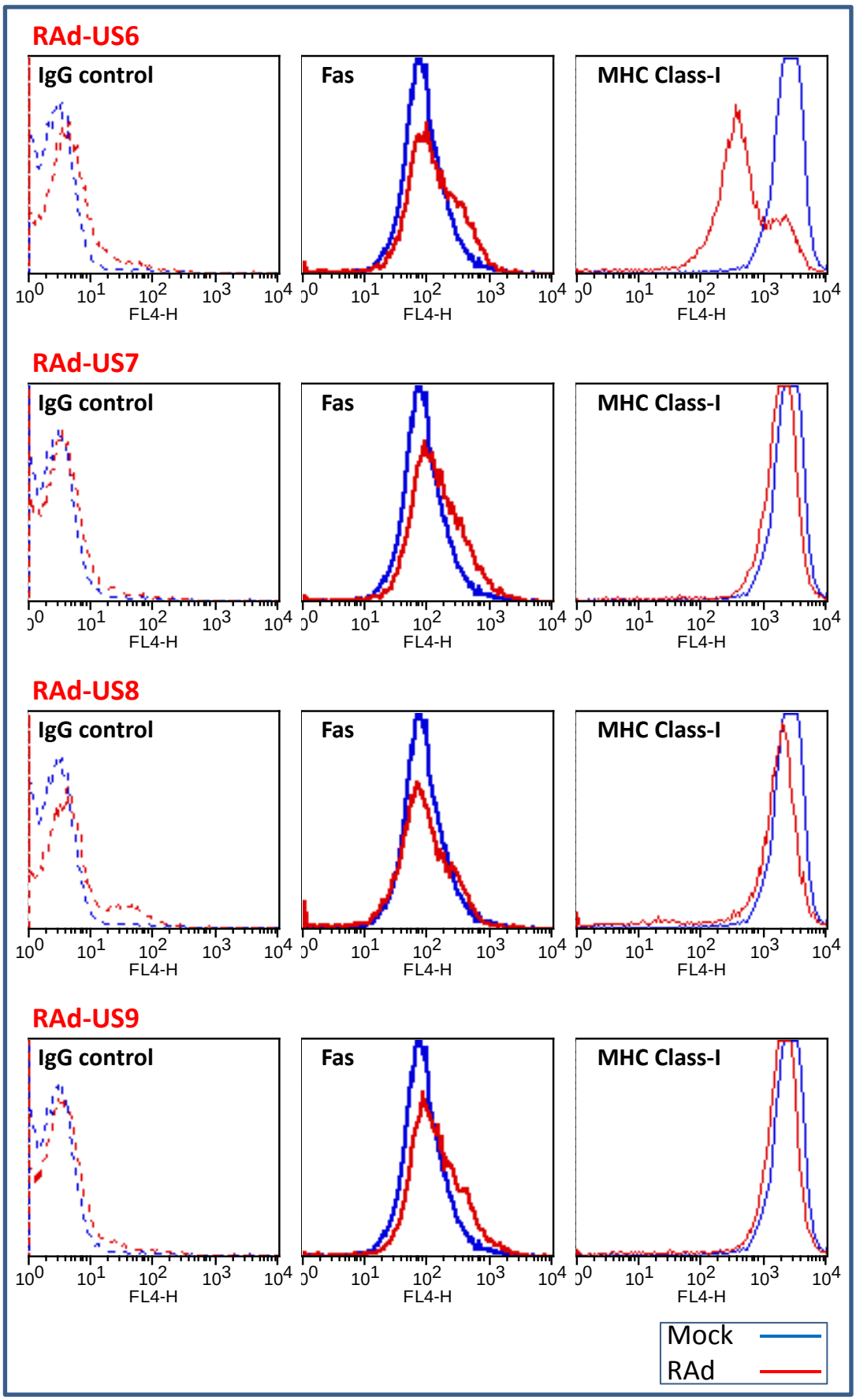
4.2.2. Gain of function screen using the RAd-library

Next, the RAd library of HCMV genes (see Chapter 3) was used in a gain-of-function assay for Fas downregulation. HFFF-hCARs were infected with Ad recombinants, each at an MOI of 10, and Fas expression detected 3 days p.i. by flow cytometry. In line with published results, RAds expressing US2, US3, US6 and US11 each downregulated MHC-class I surface expression (Figure 4.5). Otherwise levels of surface MHC class-I staining served to provide an indicator of cytotoxicity. The screen was repeated at lower MOIs for viruses that perturbed MHC-class I surface expression. In addition, a subset of Ad recombinants was associated with overt cytotoxicity during expression analysis of the RAd-library (see Section 3.4.2); these constructs were also re-screened at an MOI of 1.

None of the 170 RAds screened was observed to downregulate cell surface Fas; the complete negative dataset is not reproduced. For illustration, representative data are shown in Figure 4.5 for RAds encoding US1-US11, and the vector control. The US1-US11 block was selected to illustrate that US2, US3, US6 and US11 all clearly downregulated MHC class-I, a result that validates the use of the Ad library in functional screening.

The results of the gain-of-function assay suggest expression of no single HCMV Merlin gene is sufficient for Fas downregulation. It is possible that the function responsible was not encoded by an annotated gene, and thus was absent from the screen. Expression from 8 genes was not detected by western blot or Immunofluorescence (see Section 3.3.6). Of these, RL6, US9, US14 and US17 are deleted in block deletion mutants used in loss-of-function screening. UL146 is encoded in the UL/*b*' region, and is thus missing from strain AD169. HCMV deletion mutants covering RL9A, UL30A and UL33 were not available; nor were constructs of non-coding RNAs (except β-2.7) or miRNAs included in this assay. Moreover, it was clear that a number of HCMV genes work in combination, often as a multiprotein complex (see Sections 3.4.3.2 and 3.4.4.2). If more than one protein-coding gene was required to provide for efficient Fas downregulation, then it would not necessarily have been found by screening the Rad library.


RAd-control
IgG control
Fas
MHC Class-I
RAd-US1
IgG control
Fas
MHC Class-I
RAd-US2
IgG control
Fas
MHC Class-I
RAd-US3
IgG control
Fas
MHC Class-I
FL4-H
Mock
RAd


RAd-US6
IgG control
Fas
MHC Class-I
RAd-US7
IgG control
Fas
MHC Class-I
RAd-US8
IgG control
Fas
MHC Class-I
RAd-US9
IgG control
Fas
MHC Class-I
FL4-H
Mock
RAd

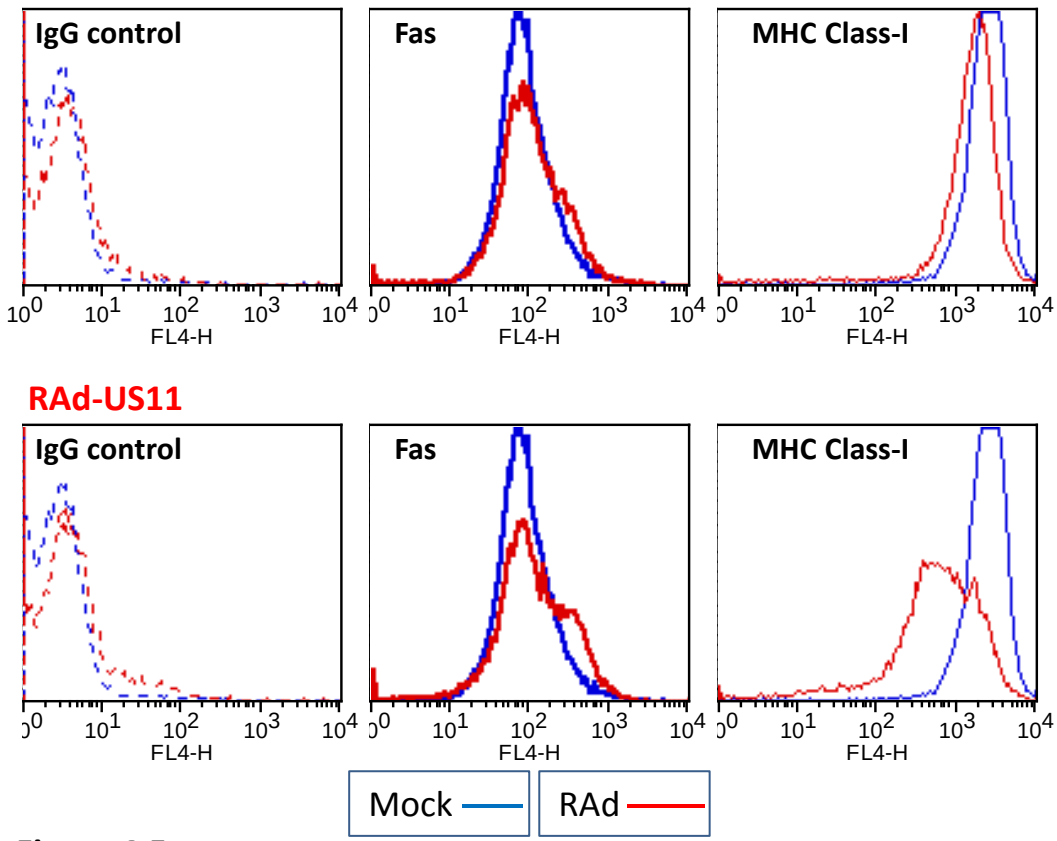


Figure 4.5

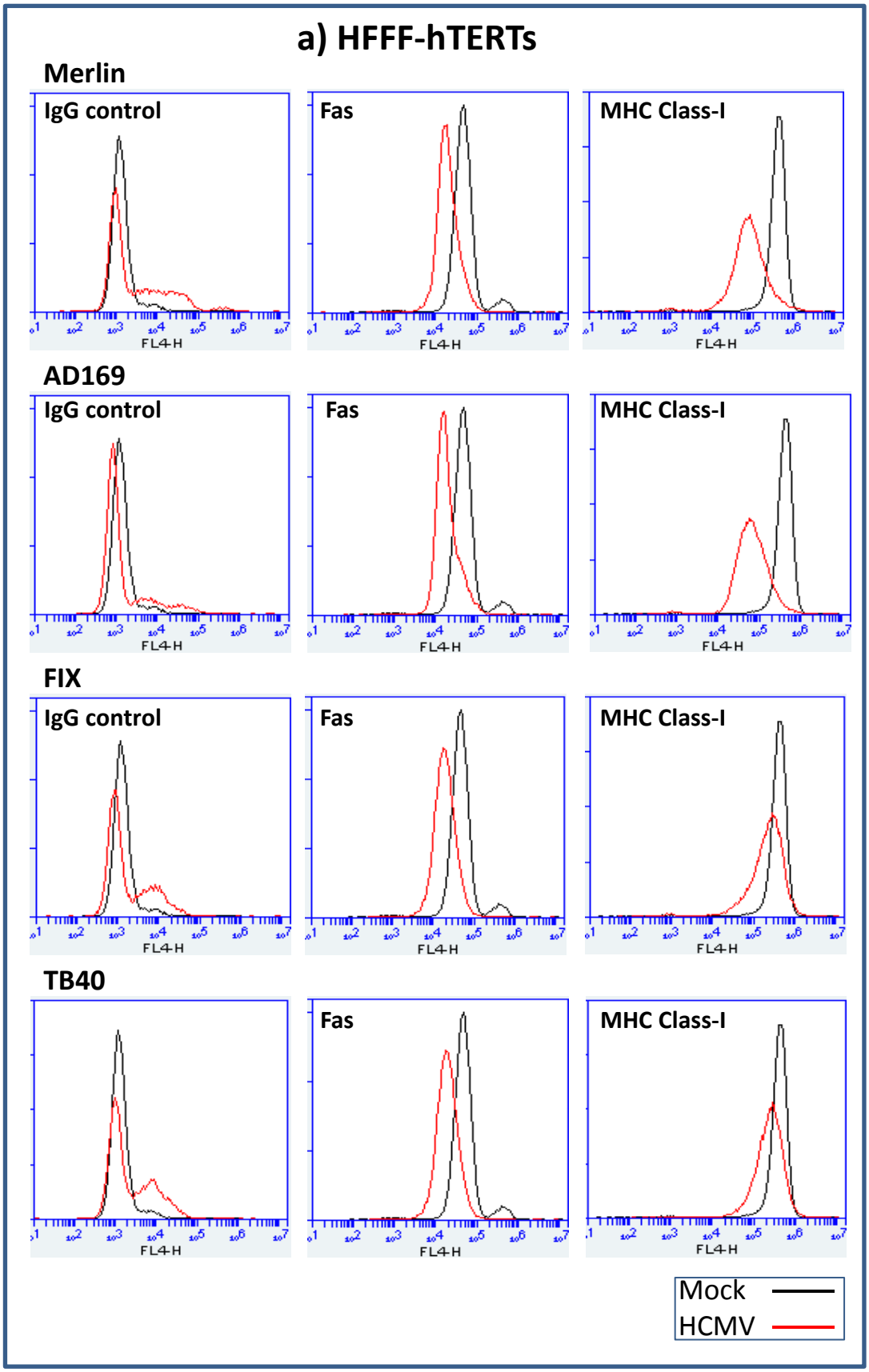
Gain of function screening of RAd-expressed HCMV genes for Fas downregulation

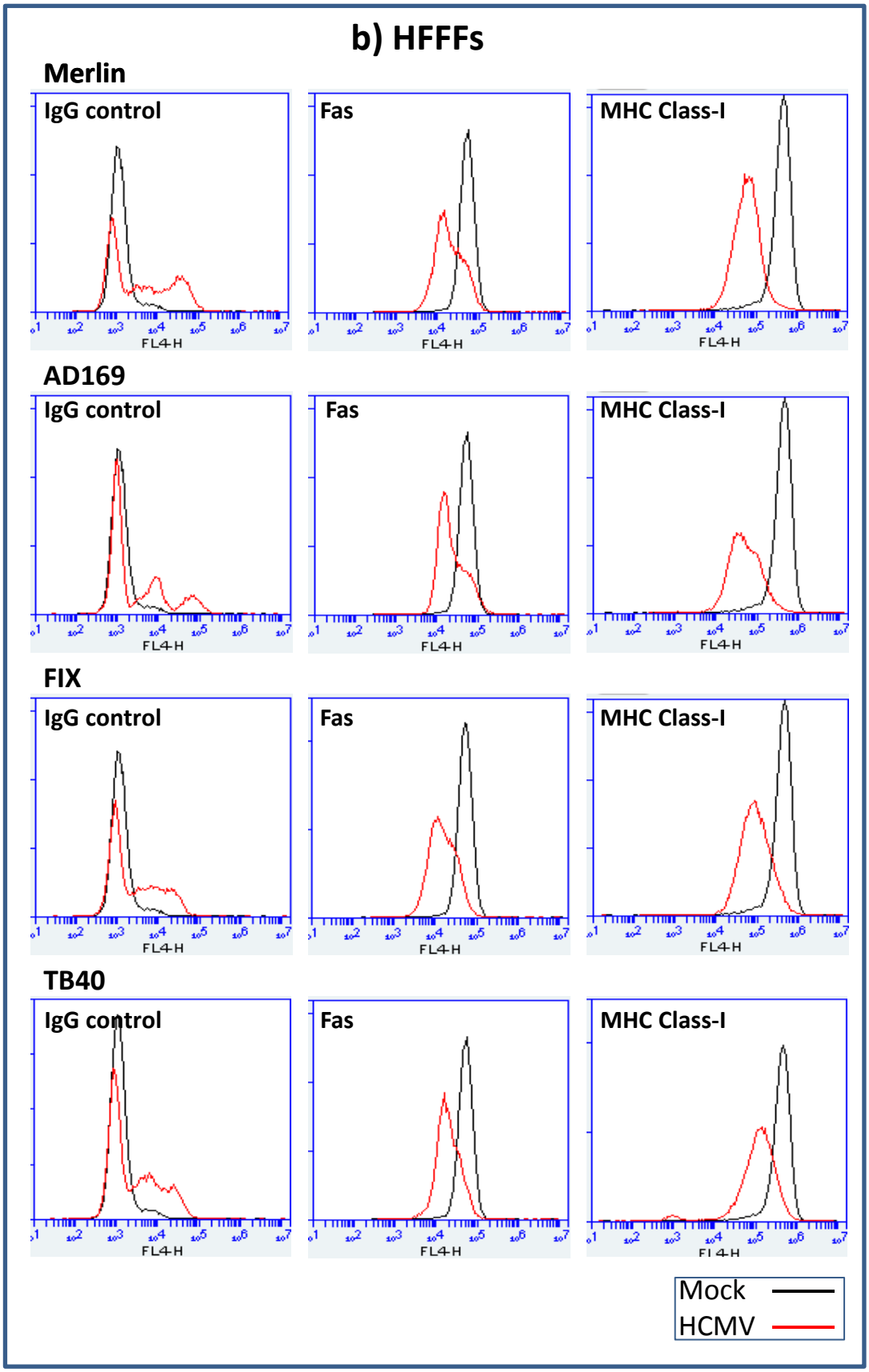
Fas expression on the surface of RAd-infected and mock-infected cells was analysed by flow cytometry in a gain of function assay. HFFF-hCARs were infected at MOI 10 (except where specified otherwise in section 4.2.2) with RAds expressing HCMV genes, or an empty control RAd. 3 days p.i., cells were stained with Fas, MHC class-I or IgG control primary antibodies, and then an Alexa fluor 647 secondary antibody. Control IgG (dotted lines), Fas (thick lines) and MHC Class-I (thin lines) staining are shown in the left-hand, middle and right-hand histograms respectively. Staining of mock samples is shown in blue, and infected cell staining displayed in red. Infection with the control RAd did not modulate Fas expression; a typical histogram is shown. Only RAds expressing genes US1 to US11 are included here.

4.2.3. Fas downregulation is a conserved function between HCMV strains

I sought to extend loss of function screening to ULb' genes, and other genes missing/inactive in other available HCMV strains. I was also interested to know if Fas downregulation is a conserved function between HCMV variants. Cell surface Fas expression was compared between cells infected with the Merlin, AD169, FIX and TB40 viruses in HFFF-hTERT, HFFF, and primary fibroblasts (donor DC) at 3 days p.i.. As shown in Figure 4.6, Fas was efficiently and equally downregulated by all 4 strains. Compared to cells infected with strains Merlin and AD169, FIX- and TB40-infected cells displayed incomplete downregulation of MHC class-I molecules from mock levels. This is explained by the fact that the BACs from which strains FIX and TB40 were derived lack some known MHC class-I-downregulating genes (FIX lacks US2, US3 and US6, whilst TB40 is missing US3 and US6.

Although it was not possible to map the gene(s) responsible for Fas downregulation, loss of function screening using a combination of Merlin block deletion mutants, and alternative HCMV strains led me to definitively discount 83 of 171 HCMV Merlin genes as not required for Fas downregulation. These include 58 genes that have been deleted in our collection of Merlin block deletion mutants, UL36, UL131A, and the 20 UL/b' genes present in Merlin but inactive/absent in AD169 (Bradley et al., 2009), IRS1 and US1 which are replaced by BAC sequences in FIX and TB40, and UL128 which is mutated in Merlin itself (Dolan et al., 2004).


a) HFFF-hTERTs
Merlin
IgG control
Fas
MHC Class-I
AD169
IgG control
Fas
MHC Class-I
FIX
IgG control
Fas
MHC Class-I
TB40
Fas
MHC Class-I
FL4-H
Mock
HCMV


b) HFFFs
Merlin
IgG control
Fas
MHC Class-I
AD169
IgG control
Fas
MHC Class-I
FIX
IgG control
Fas
MHC Class-I
TB40
IgG control
Fas
MHC Class-I
FL4-H
Mock
HCMV

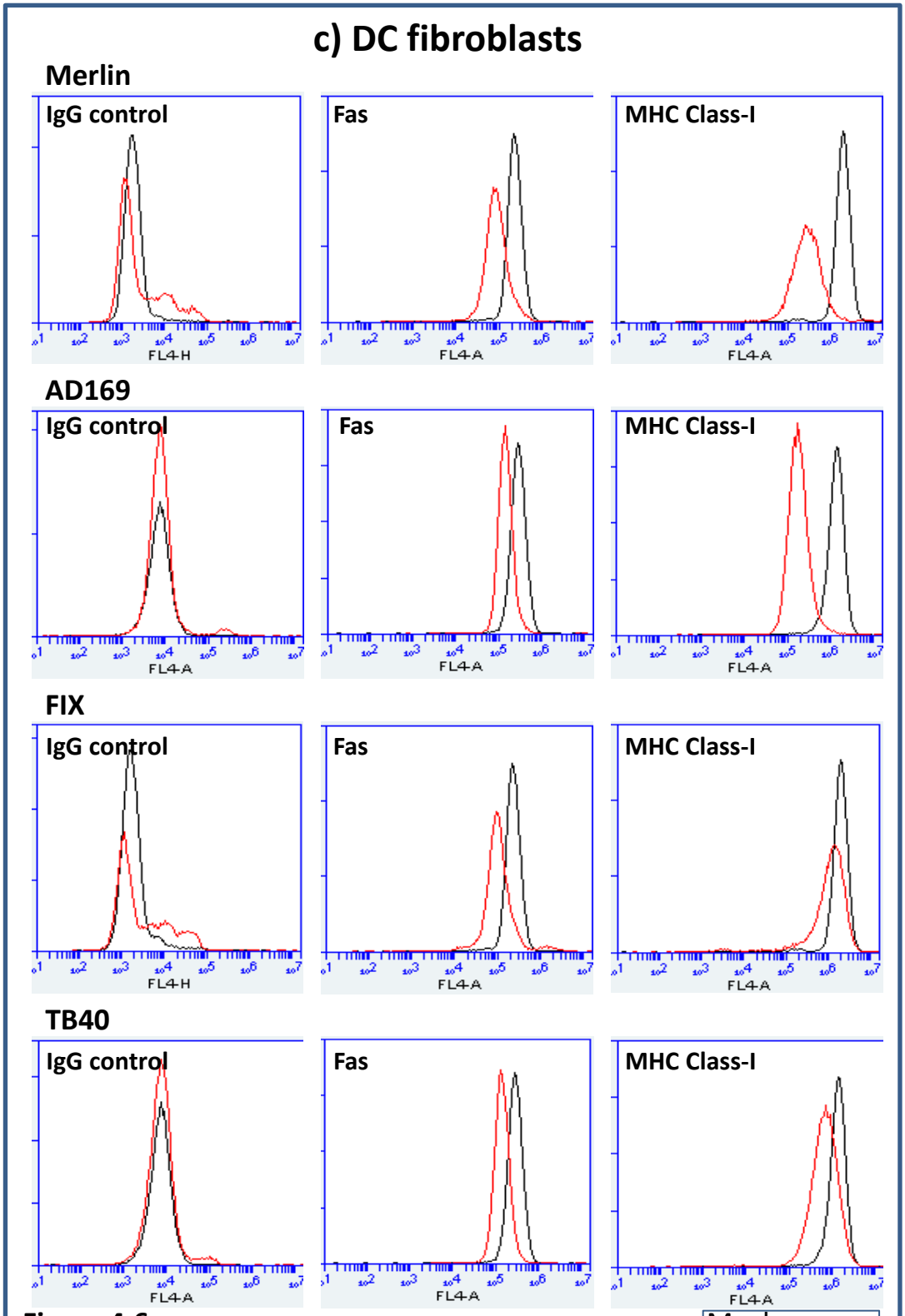


Figure 4.6

Fas downregulation is a conserved function between HCMV strains

Reduction of Fas levels on the surface of a) HFFF-hTERTs, b) HFFFs and c) primary DC fibroblasts by strains Merlin, AD169, FIX orTB40. Cells were infected at MOI 10 and harvested 3 days p.i.. Infected cells were stained with Fas, MHC class-I or IgG control primary antibodies, and an Alexa fluor 647 secondary antibody. Control IgG, Fas and MHC Class-I staining are shown in the left-hand, middle and right-hand histograms respectively. Staining of mock and infected samples are shown in black and red respectively.

4.2.4. Fas downregulation requires *de novo* viral protein synthesis

Prolonged exposure to gamma irradiation is known to cause irreparable DNA damage, and ablate *de novo* protein expression. Infection of cells with gamma-irradiated HCMV allows functions that require *de novo* viral protein expression, to be distinguished from functions performed by virion proteins. To examine whether *de novo* protein synthesis is required for Fas downregulation, strain Merlin virus was irradiated overnight with gamma rays, and its effect on cell surface Fas expression compared to that of active Merlin virus. Since gamma irradiation was performed at room temperature, it was necessary to also include infection with a separate Merlin sample that had also been kept at room-temperature. As shown in Figure 4.7, cells infected with gamma irradiated virus did not show Fas downregulation; in fact an upregulation of Fas was observed on the surface of these cells. Surface expression of MHC class-I also increased slightly, recapitulating the effect seen with the ΔUS2-US11 mutant in the loss of function assay. Cells infected with the room temperature control virus showed efficient downregulation of Fas and MHC class-I molecules, demonstrating that incubation at room temperature was not responsible for loss of Fas-downregulating function. Taken together, these data affirm that de novo HCMV gene expression and protein synthesis are required for downregulation of cell surface Fas.

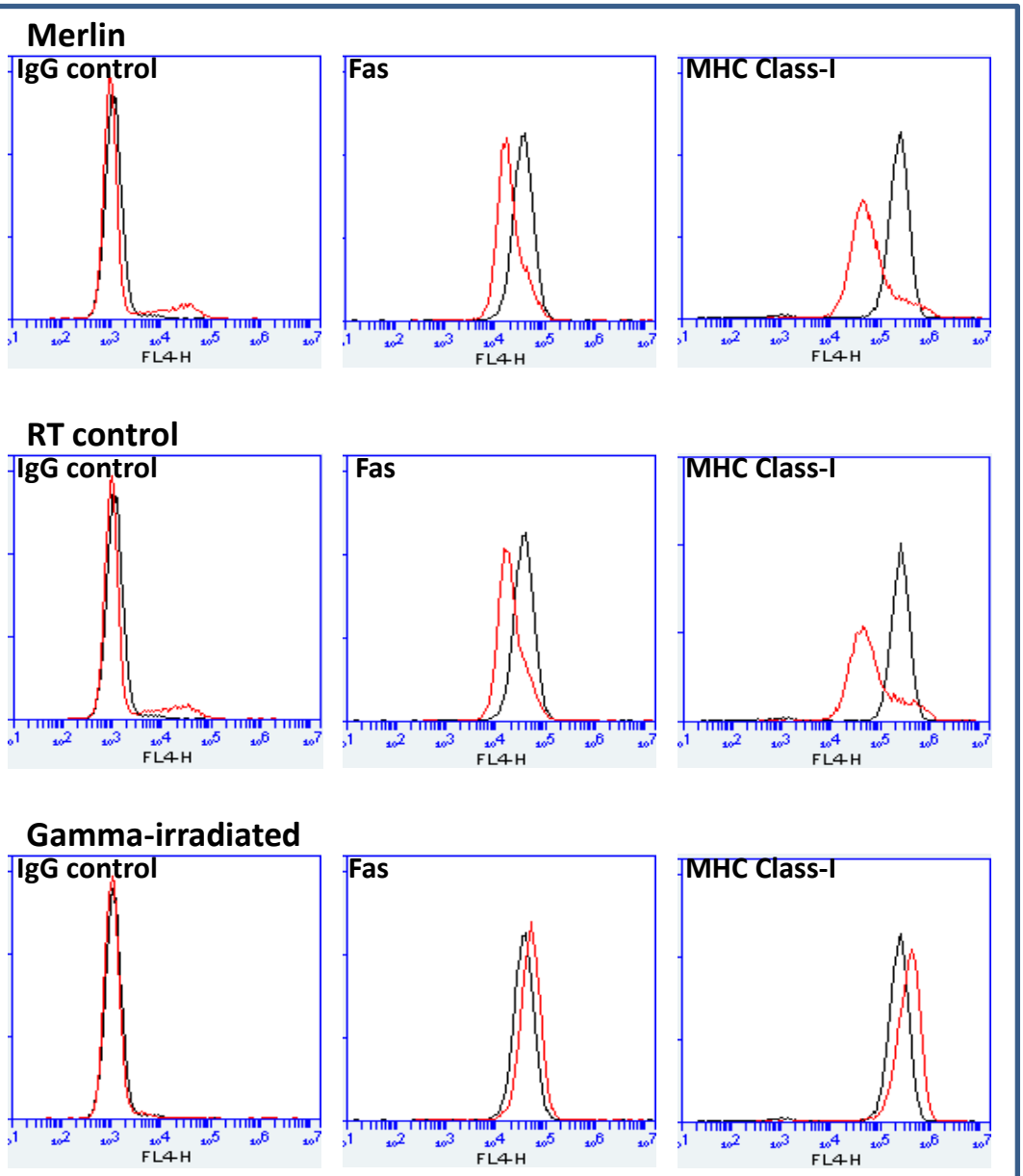


Figure 4.7

Fas downregulation by strain Merlin requires de novo viral gene expression

Surface Fas expression was compared between cells infected with Merlin, Gamma-irradiated Merlin, and a temperature control for gamma irradiation. At 3 days p.i., cells were stained with Fas, MHC class-I or IgG control primary antibodies, and an Alexa fluor 647 secondary antibody. Control IgG, Fas and MHC Class-I staining are shown in the left-hand, middle and right-hand histograms respectively. Mock and Merlin samples are shown in black and red respectively.

4.2.5. Fas downregulation begins at early times of infection

Fas downregulation was clearly evident as early as 24 hours p.i. (Figure 4.1). To characterise the temporal aspect of this function more precisely, the experiment was repeated at earlier time-points (Figure 4.8). Fas expression levels remained unchanged compared to mock-infected controls at 2, 6, 12 and 18 hours p.i., but were markedly reduced at 24 hours p.i.. Similarly, infected cells displayed impaired MHC class-I presentation only at 24 hours p.i.. This concurs with published reports describing the kinetics of MHC class I downregulation (Ahn et al., 1996, Hengel et al., 1996, Park et al., 2002).

The results of this experiment evinced that Fas downregulation begins at early stages of HCMV infection. The fact that no change in cell surface Fas expression was observed 2 hours p.i. or in cells infected with gamma-irradiated HCMV virions further suggests virus binding alone is insufficient for Fas downmodulation.

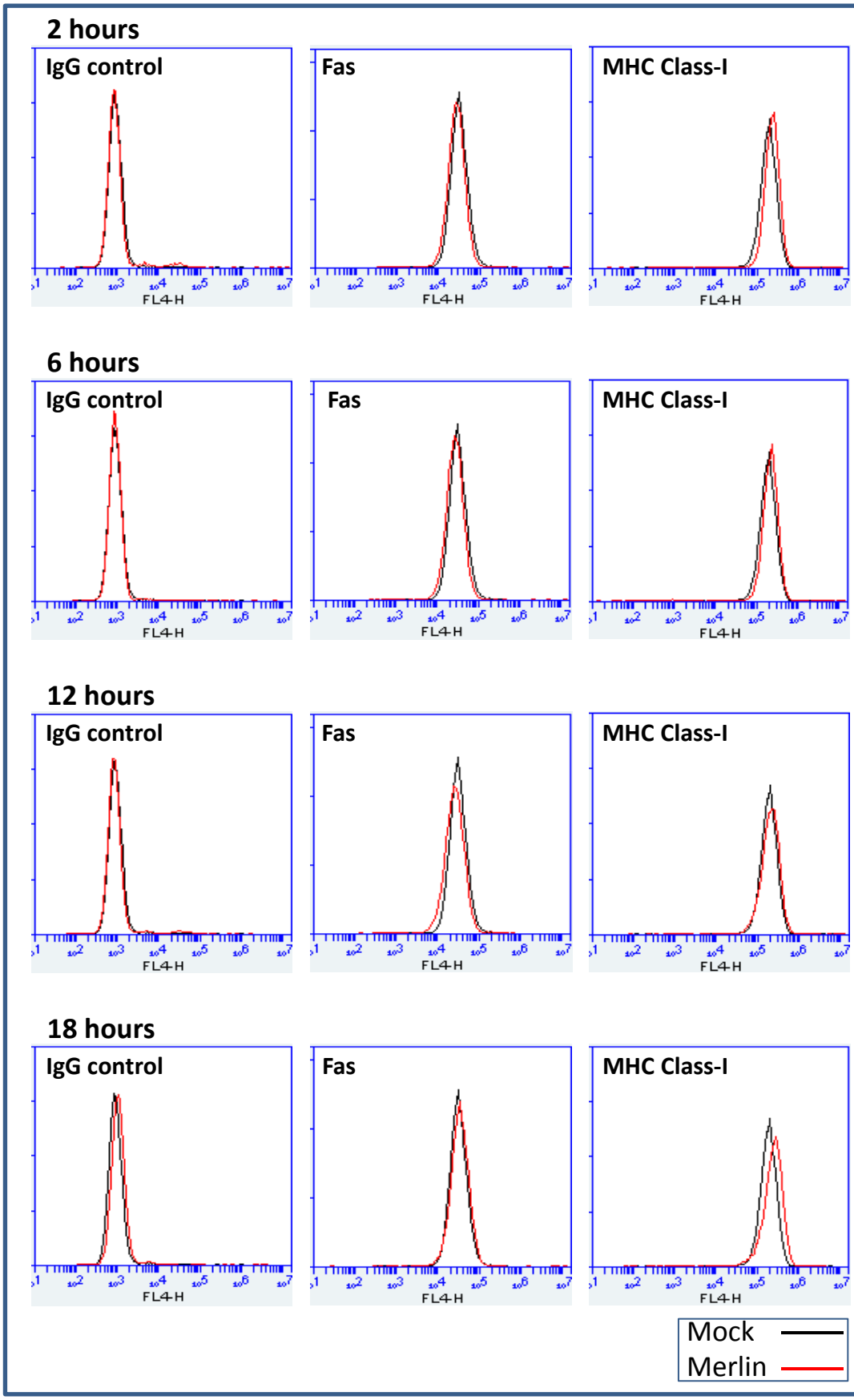
4.2.6. Fas is downregulated at the post-transcriptional level

A Q-PCR assay performed at 3 days p.i. showed Fas mRNA levels in fibroblasts were unchanged by infection with strain Merlin (Daniel Sugrue, personal communication). It was therefore hypothesised that HCMV may regulate Fas at a post-transcriptional level. To investigate whether Fas downregulation is limited to surface expression or extends to a reduction in total protein levels, Fas levels in lysates of HCMV-infected and mock-infected cells were compared. Cells infected with strains Merlin, AD169, FIX or TB40 were harvested on days 1, 2, and 3 p.i.. As illustrated in Figure 4.9a, no clear difference in Fas levels was observed between any of the 5 samples at 1 day p.i.. At 2 days p.i., Fas appeared more abundantly expressed in TB40-infected cells compared to the other 4 samples, whilst infection with strain FIX resulted in some decrease in Fas expression. On day 3, Fas levels appeared modestly diminished in cells infected with each of the 4 virus strains. While Fas expression decreased significantly at the cell surface, only a modest decrease in absolute levels of Fas in infected cells were observed.

4.2.6.1. Fas glycosylation state

Fas is a type I transmembrane receptor, and has previously been reported to be N-glycosylated (Keppler et al., 1999, Li et al., 2007, Peter et al., 1995, Shatnyeva et al., 2011). To

assess whether surface downregulation of Fas may occur as a results of ER retention, total cell lysates were treated with EndoH or PNGaseF enzymes. Proteins tend to become resistant to EndoH during maturation through the Golgi apparatus, but remain sensitive if retained in the ER. Sensitivity to EndoH in infected cells but not in mock cells would therefore indicate retention in the ER compartment. As shown in Figure 4.9b, Fas in Mock and Merlin-infected cells appeared resistant to EndoH treatment, but sensitive to PNGaseF treatment. PNGaseF treatment reduced the apparent molecular weight of Fas to 34 kDa; this is consistent with the predicted size of ~35 kDa for un-glycosylated Fas (Garcia-Fuster et al., 2004). It thus appears that HCMV-mediated Fas downregulation is not caused by ER-retention.


2 hours
IgG control
Fas
MHC Class-I
6 hours
IgG control
Fas
MHC Class-I
12 hours
IgG control
Fas
MHC Class-I
18 hours
IgG control
Fas
MHC Class-I
FL4-H
Mock
Merlin

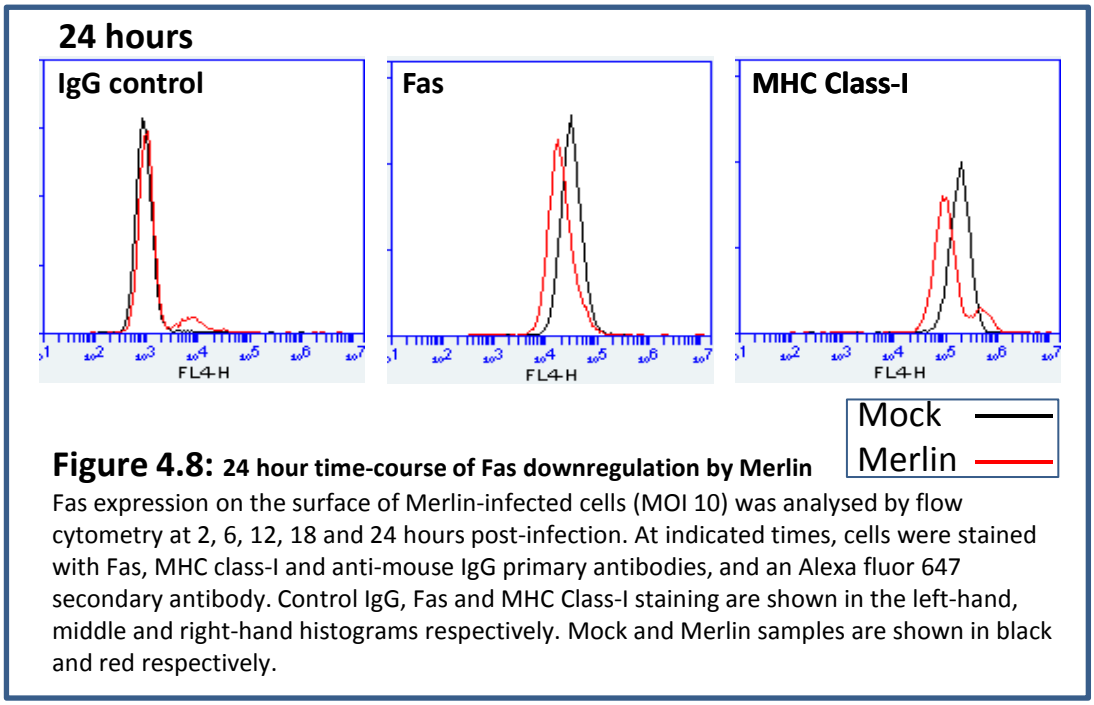


Figure 4.8: 24 hour time-course of Fas downregulation by Merlin
Fas expression on the surface of Merlin-infected cells (MOI 10) was analysed by flow cytometry at 2, 6, 12, 18 and 24 hours post-infection. At indicated times, cells were stained with Fas, MHC class-I and anti-mouse IgG primary antibodies, and an Alexa fluor 647 secondary antibody. Control IgG, Fas and MHC Class-I staining are shown in the left-hand, middle and right-hand histograms respectively. Mock and Merlin samples are shown in black and red respectively.

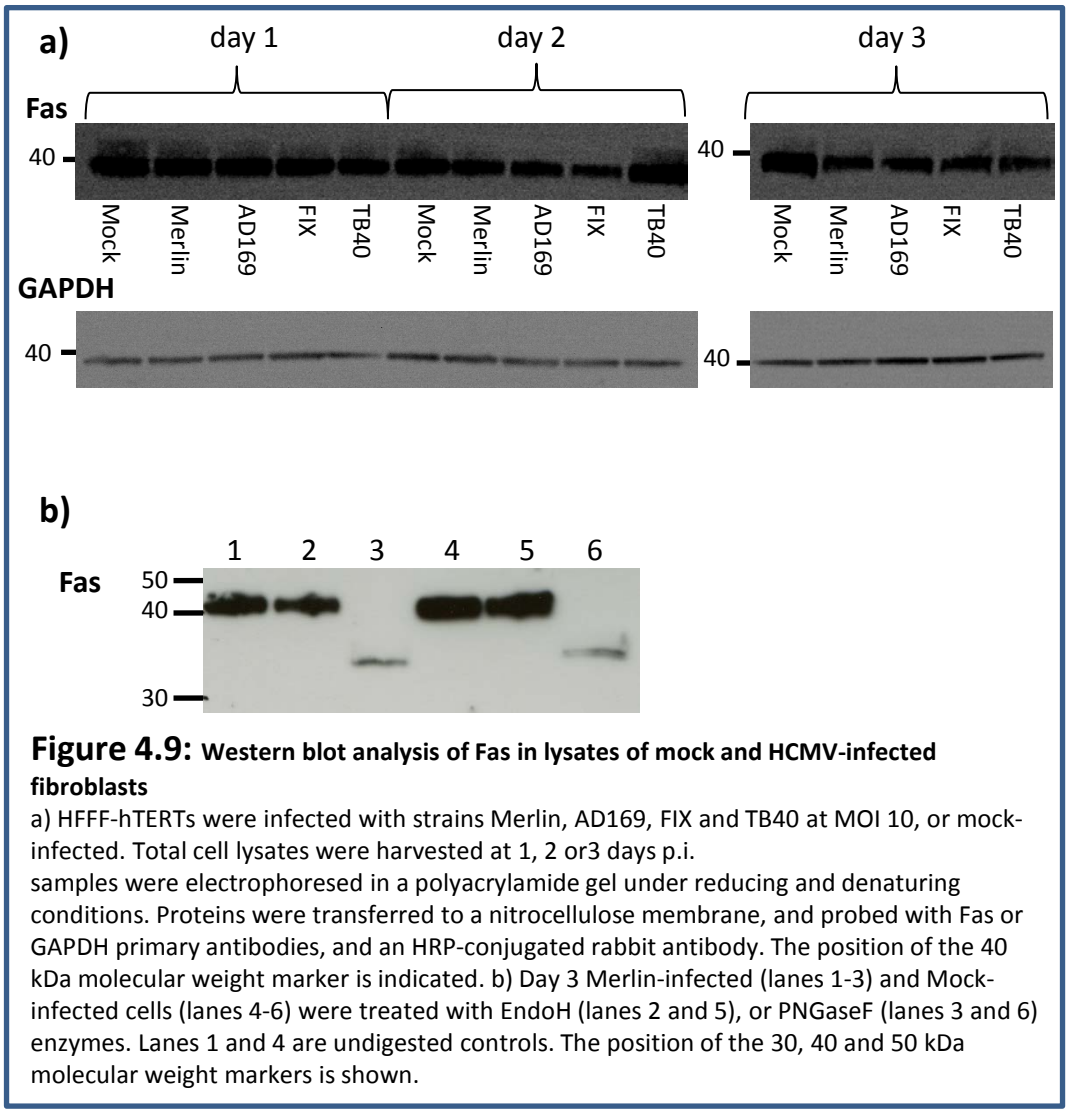


Figure 4.9: Western blot analysis of Fas in lysates of mock and HCMV-infected fibroblasts

a) HFFF-hTERTs were infected with strains Merlin, AD169, FIX and TB40 at MOI 10, or mock-infected. Total cell lysates were harvested at 1, 2 or3 days p.i.
samples were electrophoresed in a polyacrylamide gel under reducing and denaturing conditions. Proteins were transferred to a nitrocellulose membrane, and probed with Fas or GAPDH primary antibodies, and an HRP-conjugated rabbit antibody. The position of the 40 kDa molecular weight marker is indicated. b) Day 3 Merlin-infected (lanes 1-3) and Mock-infected cells (lanes 4-6) were treated with EndoH (lanes 2 and 5), or PNGaseF (lanes 3 and 6) enzymes. Lanes 1 and 4 are undigested controls. The position of the 30, 40 and 50 kDa molecular weight markers is shown.

4.2.7. HCMV infection protects against Fas-mediated apoptosis

In order to ascertain whether cell surface Fas downregulation by HCMV correlated with enhanced resistance to Fas-mediated apoptosis, the activities of effector caspases 3 and 7 following induction of Fas were compared between mock-infected and Merlin-infected cells. In order to induce Fas-signalling, cells were treated with Fas ligand or a cross-linking Fas mAb; soluble TRAIL receptor-2 or IgM control antibodies were included to control for the specificity of Fas targeting. Caspase 3/7 activity was measured by the Caspase-Glo 3/7 assay. In this assay, caspases 3 and 7 cleave a luminogenic substrate in the presence of a recombinant luciferase; this reaction results in the generation of a stable 'glow-type' luminescent signal. The amount of luminescence is proportional to the amount of caspase activity, and can be measured in a plate-reader.

Compared to mock cells, strain Merlin-infected cells were less sensitive to Fas signalling induced by either Fas ligand or Fas antibody (Figure 4.10, P=<0.001). No significant difference in caspase activity between mock and Merlin-infected cells was observed after treatment with media only, soluble TRAIL receptor-2, or IgM control antibody. The results of this assay indicate that HCMV infection elicits protection from Fas-mediated apoptosis. In this assay, the anti-apoptotic contribution of Fas downregulation cannot be distinguished from those of pUL36 and pUL37. While it is expected that reduced levels of cell surface Fas lead in turn to a reduction in Fas signalling, this could only be confirmed using a ΔUL36, ΔUL37 virus. UL37 is known to be essential for virus replication *in vitro* (Goldmacher et al., 1999), whilst contradictory reports have been published with regards dispensability of UL36 (Dunn et al., 2003c, Patterson and Shenk, 1999, Smith and Pari, 1995). Deletion of UL36 from the Merlin BAC resulted in a virus with severe growth defect (results not shown). The relative contributions of Fas down-regulation, and pUL36 and pUL37 activity towards inhibition of Fas-mediated apoptosis could therefore not be dissected in the context of HCMV infection.

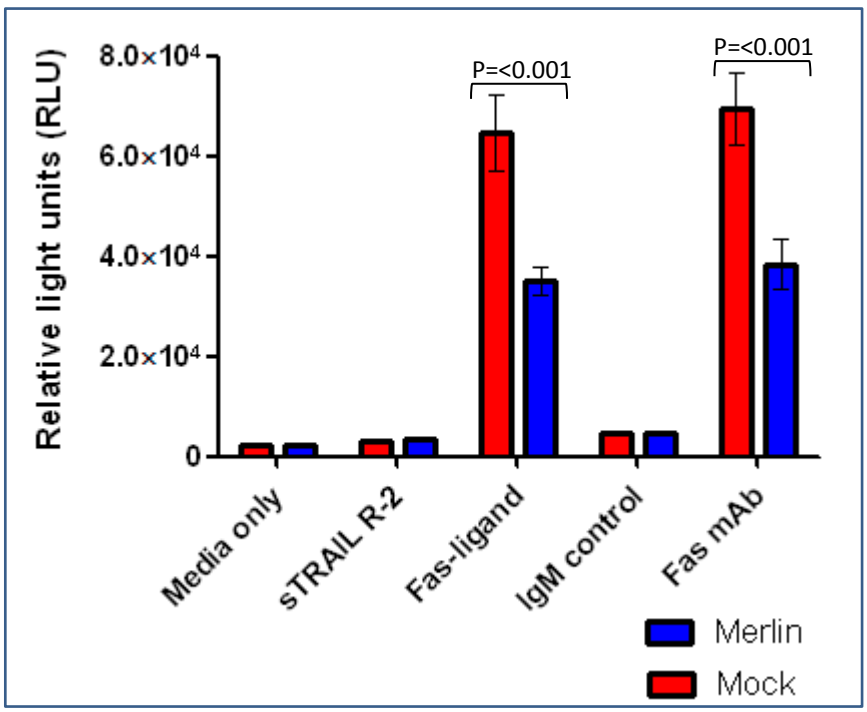


Figure 4.10: Strain Merlin-infected cells are more resistant to Fas-mediated apoptosis than un-infected controls

$5x10^3$ HFFF-hTERTs were infected with Merlin at MOI 10, or mock-infected. At 60 hours p.i., cells were treated with soluble TRAIL receptor-2, Fas ligand, IgM isotype control antibody, or Fas mAb. At 72 hours p.i., cells were incubated with Caspase-Glo 3/7 reagent for 1 hour. Luminescence generated by caspase 3/7 activity was measured in relative light units (RLU) in a plate reader. Results are presented as means ± SEM of quadruplicate samples.

4.3. Summary

In short, it was demonstrated that infection of fibroblasts with HCMV results in reduction of cell surface Fas expression. This phenotype was first detected at 24 hours p.i., and persisted into true-late stages of infection. Moreover, reduced Fas cell surface expression correlated with resistance to Fas-mediated anti-apoptotic activity. Loss-of-function and gain-of-function screens using available HCMV deletion mutants and RAd-encoded HCMV genes respectively could not identify the function(s) responsible. Fas was preferentially removed from the cell surface while intracellular stores appeared relatively intact. This argues protein degradation is unlikely to be responsible for decreased expression of Fas on the cell surface. In addition, the fact that Fas is resistant to EndoH digestion in both un-infected and Merlin infected cells suggests ER-retention does not account for HCMV-mediated Fas downregulation. The possible mechanism of Fas downregulation is further discussed in chapter 7

5- Characterisation of the novel HCMV UL150A gene

5.1. Alternative splicing in HCMV

Alternative splicing is the process by which a single primary RNA transcript (pre-mRNA) is differentially spliced to produce multiple mRNA variants. As a mechanism of generating functional diversity, alternative splicing has a number of advantages, such as low genetic expenditure. Alternative splicing is particularly useful in the case of viruses such as HCMV, where genome packaging is under strict size constraint.

Nuclease mapping first identified the MIE pre-mRNA to be a spliced transcript containing 4 exons, separated by 3 introns (Stenberg et al., 1984, Akrigg et al., 1985, Wilkinson et al., 1984). At least 5 different mRNA molecules are known to arise from the MIE pre-mRNA, including transcripts for IE72 and IE86 (Akrigg et al., 1985, Awasthi et al., 2004, Kerry et al., 1995, Shirakata et al., 2002). Expression of the US3 gene produces an unspliced mRNA, and 2 additional spliced variants (Tenney et al., 1993, Weston, 1988). By nuclease mapping, 3 alternatively spliced transcripts were found to originate from the UL37 gene (Kouzarides et al., 1988). Subsequent studies on the UL37 gene discovered several additional spliced transcripts, many of which are expressed at extremely low amounts during HCMV infection, but whose protein products have not been detected (Adair et al., 2003, Goldmacher et al., 1999, Su et al., 2003). The UL112 transcript undergoes complex splicing; truncated 2.1 kb and 2.2 kb transcripts were abundant at early times of infection, but gradually replaced by a larger splice variant of 2.5 kb, as well as the unspliced 2.65 kb transcript at late times p.i. (Wright and Spector, 1989). Leatham et al. reported that transcription from UL119-UL115 gives rise to alternatively spliced transcripts at different times of infection (Leatham et al., 1991). Later, RT-PCR of cDNA from HCMV-infected cells using primers designed to bind splice sites revealed novel splice variants in regions corresponding to the UL22A, UL89, RNA5.0 and UL119 genes (Rawlinson and Barrell, 1993). The UL33 gene was also found to undergo splicing (Margulies et al., 1996), and more recently, characterisation of RNA5.0 demonstrated that splicing of its 2 exons generates RNA1.1 (Kulesza and Shenk, 2004). Kotenko et al. demonstrated that the viral IL-10 mRNA was formed by the splicing of 2 exons, separated by an intron present in the primary transcript (Kotenko et al., 2000b). The spliced UL128 and UL131A genes were predicted from sequence comparisons between HCMV and CCMV, and their splicing patterns confirmed by mRNA mapping (Akter et al., 2003). Discovery of a polypeptide derived from both the UL28 and UL29 proteins prompted the hypothesis that these proteins are expressed from a spliced mRNA (Moorman et al., 2008). Subsequently it was shown that the UL28/UL29 protein is expressed from a primary transcript carrying multiple donor and acceptor sites (Mitchell et al., 2009).

5.1.1. Novel HCMV spliced transcripts revealed by deep sequencing

In an attempt to better understand HCMV gene expression during productive infection, and with particular focus on splicing, Gatherer et al. carried out deep sequencing of viral transcripts in human fibroblasts 3 days after they were infected with the Merlin strain (Gatherer et al., 2011). Analysis of transcripts revealed splicing is more common than previously anticipated; following exclusion of probable artefactual reads, a total of 389 potential splice sites including 229 donor and 132 acceptor sites were identified. Reverse Transcription-PCR (RT-PCR) and rapid amplification of cDNA ends (RACE)-analysis strongly supported these findings.

4 new genes (RL8A, RL9A, UL150A, and US33A) were identified based on information obtained from the analysis. Additionally, 5 acceptor sites were found to be spliced from a large number of upstream donor sites including exons which may 5' extend the coding region, and these are henceforth referred to as superacceptor site (SASs). SASs are located upstream of RL8A, UL74A, UL92, UL124, and UL150A (pair of SASs) predicted coding regions, and except one of the pair, all are conserved in CCMV. The SASs in the vicinity of RL8A, UL74A, UL124, and UL150A were further characterised. Figure 5.1 illustrates the position of these SASs and their predicted donor sites in the Merlin genome.

Of the 5 coding regions present in the vicinity of SASs, UL92 and UL124 are recognised ORFs. RL8A is conserved in CCMV, and its predicted protein product contains potential transmembrane regions. 5' ends of 4 upstream exons were found to be spliced to the RL8A SAS, but RL8A appeared weakly expressed, since transcript sizes generated in RACE analysis were not detected by northern blotting. UL74A is predicted to encode a membrane glycoprotein (Scalzo et al., 2009), whilst UL150A is antisense to UL150. The UL150A coding region is made up of two exons separated by an intron, which gives rise to miR-UL148D. Furthermore the 2 UL150A exons and their intervening intron are all conserved in CCMV. UL74A and UL150A SASs are spliced from the 5' ends of 4 upstream exons each, and the locations of these exons are consistent with the sizes of transcripts detected by northern blot.

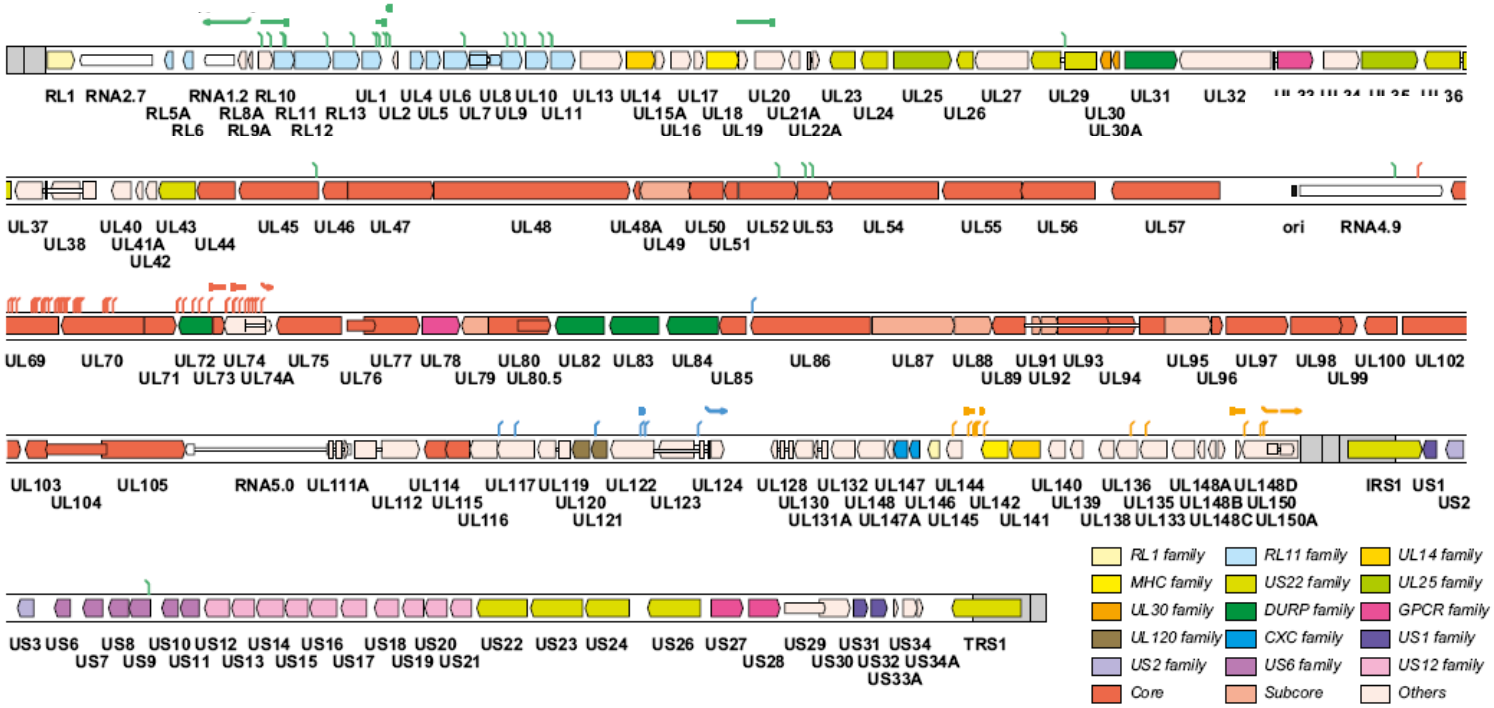


Figure 5.1 **(adapted from figure S3 of Gatherer et. al. 2011): Splicing of HCMV trancripts identified by deep sequencing**

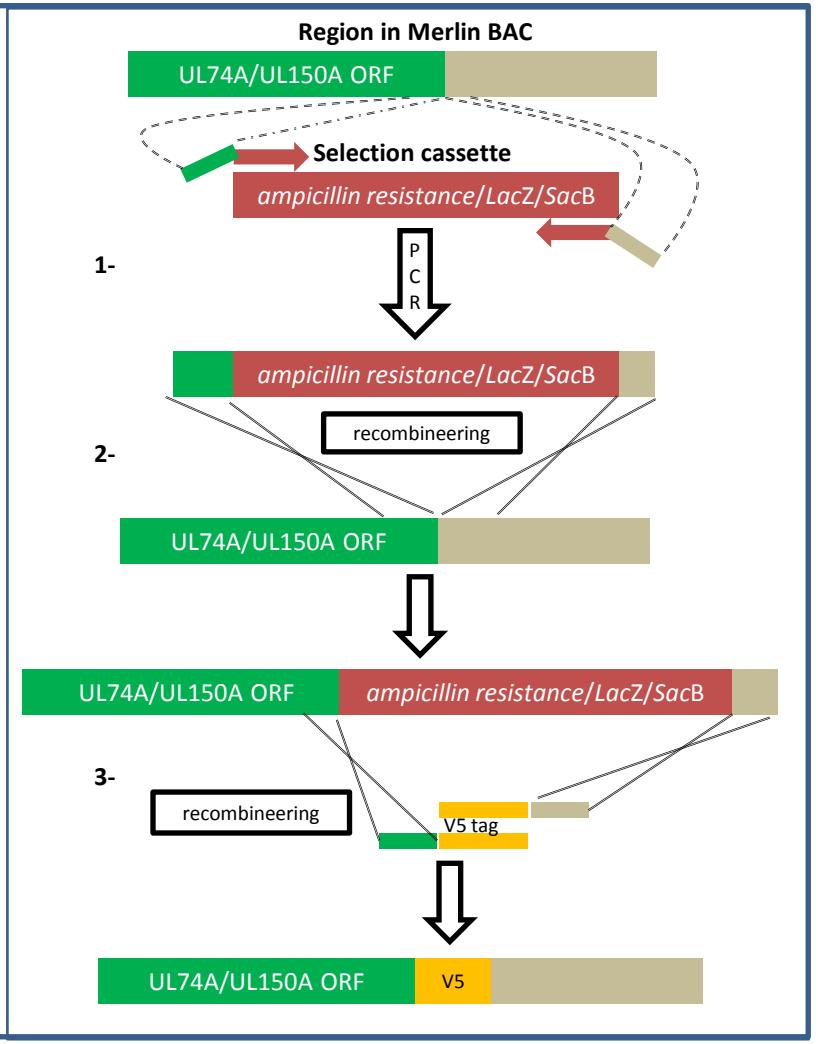
Deep sequencing the Merlin transcriptome identified 4 novel splice superacceptor sites (SASs) from RNA splicing. 3' exons of RL8A (green), UL74A (red), UL124 (blue) and UL150A (orange) are shown as coloured arrows with diagonal lines at the beginning. The gap between UL150A exons represents intronic sequences. An SAS was identified closely upstream of each of these 3' exons. Exons whose 5' ends were found by RACE to be spliced to SASs are shown as arrows beginning with vertical bars. Where in close proximity to each other, these upstream exons have been combined. Other donor sites that were located upstream and in the same orientation as their corresponding SAS are shown as vertical hockey stick symbols.

5.2. Generation of recombinant Merlin virus encoding tagged UL74A and UL150A genes

To validate the prediction that UL74A and UL150A are protein coding genes, and to test whether alternative splicing of their transcripts may result in the generation of diverse protein products, I analysed the expression of these genes in the context of HCMV infection. We had previously been successful in the detection of V5-tagged proteins in the Ad-library (see Chapter 3). A two-step recombineering approach was used to insert V5 epitope tags at the C-terminus of the UL74A and UL150A ORFs within the Merlin BAC (separate viruses), and the recombinant viruses were designated RCMV1569 and RCMV1571 respectively. In the first recombineering round, the *ampicillin resistance/Lac*Z/*Sac*B cassette was inserted after the UL74A and UL150A genes, and colonies containing the cassette identified by positive selection (see Section 2.3.8.1). In the second recombineering round, the cassette of selectable markers was replaced with a V5 tag and recombinants identified by negatively selecting against the cassette. Figure 5.2 contains a diagrammatic representation of the cloning steps.

Since sucrose was used in negative selection to inhibit the growth of *Sac*B-expressing bacteria, it was necessary to validate *Sac*B expression after the first recombineering round. This was done by assessing the growth of bacteria on media with or without sucrose (Figure 5.3). Colonies with mutated *Sac*B genes grew equally well on sucrose and non-sucrose plates, and were discarded. However, growth of colonies with fully functioning *Sac*B genes was inhibited by sucrose; these were selected for the second recombineering round. The sequence of regions undergoing recombineering was then validated, and recombinant HCMV generated (see Sections 2.9.1-2.9.3).

Figure 5.2: V5-tagging UL74A and UL150A genes
V5 tags were inserted at the C-terminus of the predicted UL74A (nucleotide position 109050) and UL150A (nucleotide position 194124) genes immediately before their respective stop codons. 1- The *ampicillin resistance*/*Lac*Z/*Sac*B (see Section 2.3.8.1) was amplified by PCR from the AdZ vector (see chapter 3). Primers used to amplify the cassette carried 5' overhangs homologous to regions either side of the insertion site. 2- The cassette was inserted at the C-terminus of the UL74A and UL150A genes in the first recombineering round, and recombinants identified by positive selection for the cassette. 3- In the second recombineering round, overlapping oligonucleotides replaced the cassette of selectable markers with a V5-tag. These oligos contained the V5-tag at their overlapping 3' ends, and were homologous at their 5' ends to regions upstream and downstream of the insertion site. Recombinants were identified by negative selection against the cassette.The sequence of regions undergoing recombination was validated before generation of recombinant virus. The primer sequences used for inserting tags into the Merlin BAC, and for validating recombinants are provided in Appendix II.



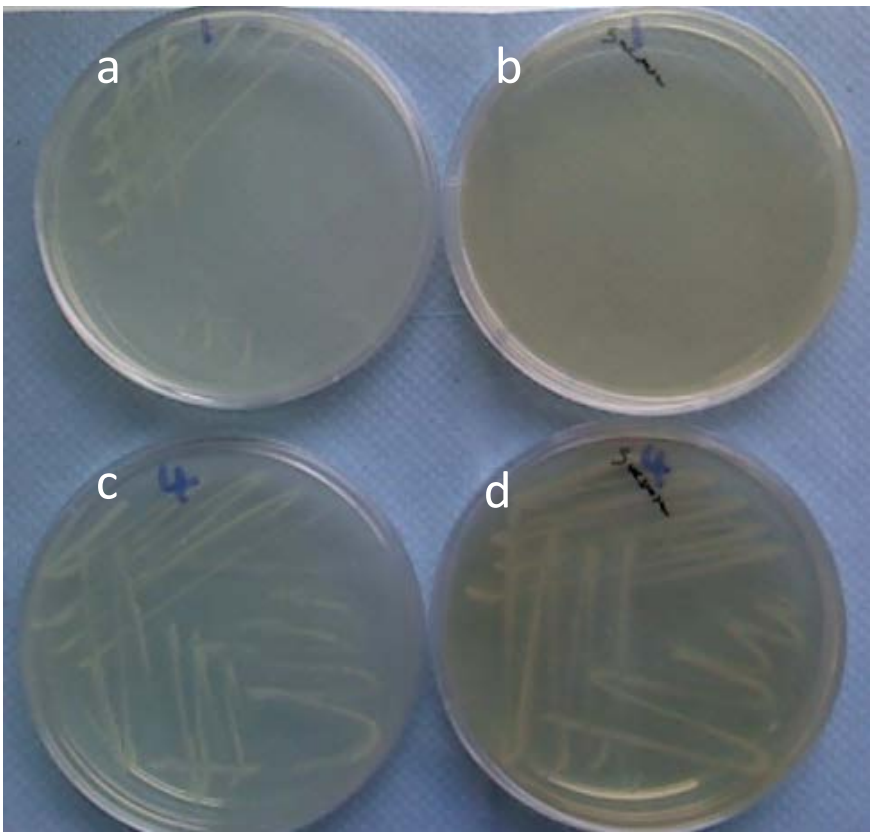


Figure 5.3:

Testing *Sac*B expression in bacteria containing the selection cassette

By positive selection, clones 1 and 4 were identified to contain the *ampicillin resistance*/*Lac*Z/*Sac*B/ cassette. Resuspensions of each colony were streaked onto +sucrose (b & d), or -sucrose (a & c) media, and plates incubated at 32°C overnight. Sucrose sensitivity is observed on plate (b) but not plate (d). Clones 1 and 4 therefore contain functioning and mutated *Sac*B respectively.

5.3. Expression analysis of UL74A and UL150A gene products

To characterise protein expression from UL74A and UL150A, HFFF cells were infected with the tagged viruses, and immunoflourescence and western blot analyses performed over a time-course. Using both methods, protein expression was confirmed for UL150A (Gatherer et al., 2011) but not UL74A (Figure 5.4). For UL150A, a 34kDa doublet was detected 1 day p.i., which became more abundant at day 3, and yet more abundant at day 6 (figure 5.4a). Another less pronounced also 50 kDa band appeared 3 days p.i.. Particularly when the actin protein loading control is considered, this higher molecular weight isoform is more abundantly expressed late in infection. The predicted mass of the tagged primary translated protein was 31 kDa, similar to that of the most abundant doublet detected. Multiple forms of pUL150A are therefore expressed during the early times of infection that become abundant as the infection progresses; a result that is consistent with the gene being differentially spliced.

pUL150A was also detected by immunofluorescence at all three time-points (figure 5.4b). It resided in close proximity to the nuclei of infected cells, often forming aggregates resembling the ER/Golgi complex.

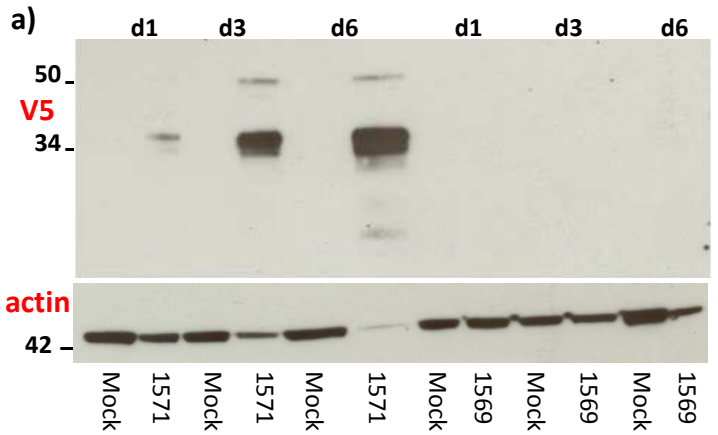
Although epitope-tagged UL74A was not detected by western blot and immunofluoresence, this protein may still be expressed by HCMV. Lack of detection could be caused by the cleavage of the C-terminal tag, or folding of the protein in a manner preventative of antibody binding.

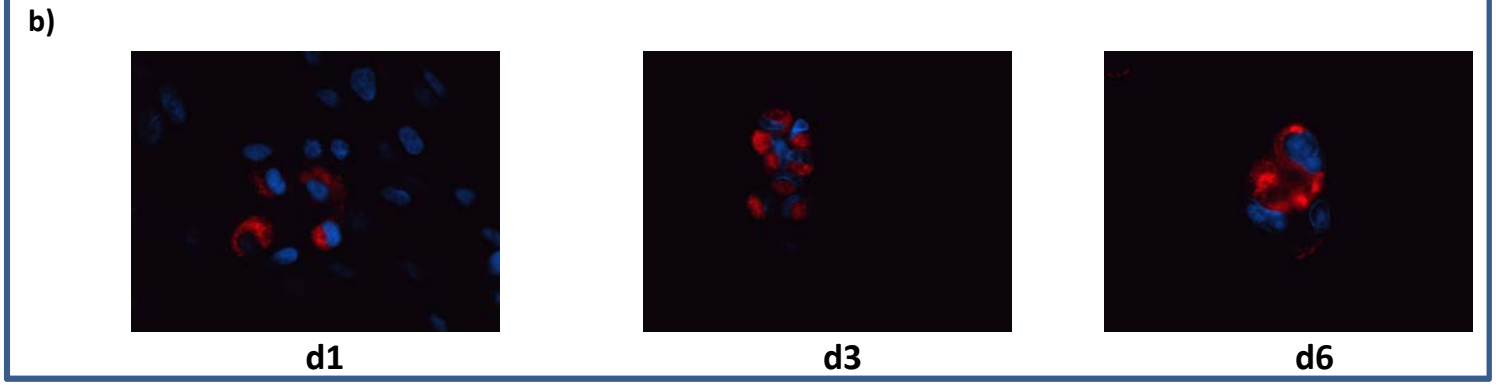
Summary

Detection of UL150A protein products in the context of HCMV infection validates the prediction of UL150A as a coding region. The main UL150A protein-coding sequence is captained within the larger UL150 gene that is present on the opposing DNA strand and transcribed in the opposite orientation; yet its amino acid sequence is better conserved than UL150. The discovery of this novel gene demonstrates for the first time the presence of overlapping coding genes on opposite strands of the HCMV genomic DNA

Figure 5.4: Timecourse expression analysis of Merlin UL74A and UL150A proteins

UL74A and UL150A genes were V5-tagged in Merlin to produce recombinant viruses RCMV1569 and RCMV1571 respectively. HFFFs were infected at MOI=5, and expression of tagged proteins analysed at 1, 3, and 6 days p.i. by a) Western blot: membrane was probed with V5 or actin primary, and an HRP-conjugated secondary antibody; b) immunofluorescence: cells were fixed and permeabilised, and HCMV proteins detected using the V5 mAb, and an Alexa Fluor 594 secondary antibody. Nuclei were stained with DAPI and expression was visualised microscopically using a x40 oil objective. No V5 expression was detected in RCMV1569-infected cells; only RCMV1571-infected cell images are shown.





6- Characterisation of gpUL4

6.1. HCMV secretome

HCMV has been shown to profoundly affect the secretome of infected cells (Dumortier et al., 2008); this mostly impacts on release of immune-suppressant proinflammatory cytokines and chemokines. Indeed, a small number of HCMV-encoded proteins have also been reported to be secreted from infected cells: (i) US22 is a nuclear protein that was also detected in infected-cell supernatants (Mocarski et al., 1988). (ii) The glycoprotein product of the UL22A gene, originally designated R27080, is a secreted chemokine receptor that selectively binds RANTES (Müllberg et al., 1999, Wang et al., 2004). (iii) gpUL111A (cmvIL-10) is able to bind the IL-10 receptor in competition with cellular IL-10 (Kotenko et al., 2000a). (iv) UL146 and UL147 proteins exhibit limited homology to the CXC chemokine IL-8, and are named vCXC-1 and vCXC-2 respectively (Penfold et al., 1999). vCXC-1 is secreted during late times of infection and binds the CXCR2 receptor with affinity comparable to IL-8. (v) More recently, UL7 has been shown to encode a type-I transmembrane glycoprotein with structural homology to CD229, a cell surface receptor of the signalling lymphocyte-activation molecule (SLAM) family. The ectodomain of gpUL7 is shed from the surface of infected cells and binds monocyte-derived dendritic cells (DCs) to modulates cytokine production (Engel et al., 2011). I wished to test whether the RAd library could be used to identify and characterise novel HCMV secreted proteins

6.2. Screening for secreted HCMV proteins

As an extension of the expression studies described in chapter 3, supernatants of cells infected with the RAd library of HCMV genes were screened by western blot using the V5 antibody. Secreted UL22A was not detected using the V5 antibody, but a polyclonal antibody raised against RAd-UL22A detected a protein closely resembling in size the reported 45kDa R207080 (data not shown). Similarly, UL7 was detected by the V5 antibody in the extracts but not the supernatants of infected cells (see Figure 3.3, supernatant data not shown). This is in agreement with the finding that the UL7 C-terminus remains intracellular following shedding of the ectodomain (Engel et al., 2011). Secreted products of UL111A, UL146, and US22 were not detected; these proteins may still be secreted at low amounts when expressed using Ad recombinants. The secreted product of one other gene (UL4) was also detected in supernatants of RAd-infected cells.

6.2.1. Secretion of gpUL4

gpUL4 was identified as being secreted during the screen of the RAd library. On a western blot, secreted gpUL4 migrated as a diffuse band corresponding to an apparent molecular mass ranging between 52-60 kDa; in comparison with a ~50 kDa species detected in extracts of RAd-UL4-infected cells (Figure 6.1a). The size difference between the intracellular and secreted forms, characteristic of secreted proteins, implied that an additional post-translational maturation step was associated with secretion, and the extracellular form did not represent contamination arising from spontaneous lysis of RAd-UL4-infected cells.

Chang et al. identified a 48 kDa product which they designated gp48 (herein referred to as gpUL4) in extracts of infected with HCMV strain Towne, using polyclonal sera raised against an *in vitro*-translated UL4 polypeptide (Chang et al., 1989). The *in vitro* synthesised polypeptide migrated at 17 kDa, consistent with the size predicted for the UL4 primary translation product. Treatment of strain Towne-infected cells with tunicamycin, an inhibitor of N-linked glycosylation, resulted in accumulation of a smaller 27 kDa protein. This was still 10 kDa larger than the primary UL4 translation product, leading the authors to suggest that gpUL4 is possibly both N- and O-glycosylated (Chang et al., 1989).

To characterise glycan modifications of gpUL4 and to test the hypothesis that secreted gpUL4 is differentially modified, lysates and supernatants of RAd-UL4-infected cells were treated with EndoH and PNGaseF enzymes (Figure 6.1b). Cell-associated gpUL4 was almost completely sensitive to EndoH treatment, appearing as a 17 kDa species. Treatment with PNGaseF reduced intracellular gpUL4 to 16 kDa, which corresponds closely with the 17 kDa predicted size of the primary translation product. Taken together these data are consistent with intracellular gpUL4 being subjected to only N-linked glycosylation and not being fully processed in the Golgi apparatus. In contrast to cell-associated gpUL4, secreted gpUL4 was largely resistant to EndoH, observed as an exceptionally diffuse ‘smear’ ranging between 18-60 kDa. PNGaseF digestion transformed secreted gpUL4 to a product migrating at 17-20 kDa, indicating O-linked glycosylation. The difference in sensitivity to EndoH and PNGaseF displayed by the two forms of gpUL4 further supported the conclusion that gpUL4 was actively secreted from RAd-infected cells. Indeed it is well known that glycosylation plays an important role in dictating appropriate transport and targeting of proteins that are destined to be secreted.

Immunofluorescence staining of RAd-infected cells showed that the intracellular gpUL4 was cytoplasmic, in some cells preferentially accumulating in structures proximal to the nucleus (Figure 6.1c).

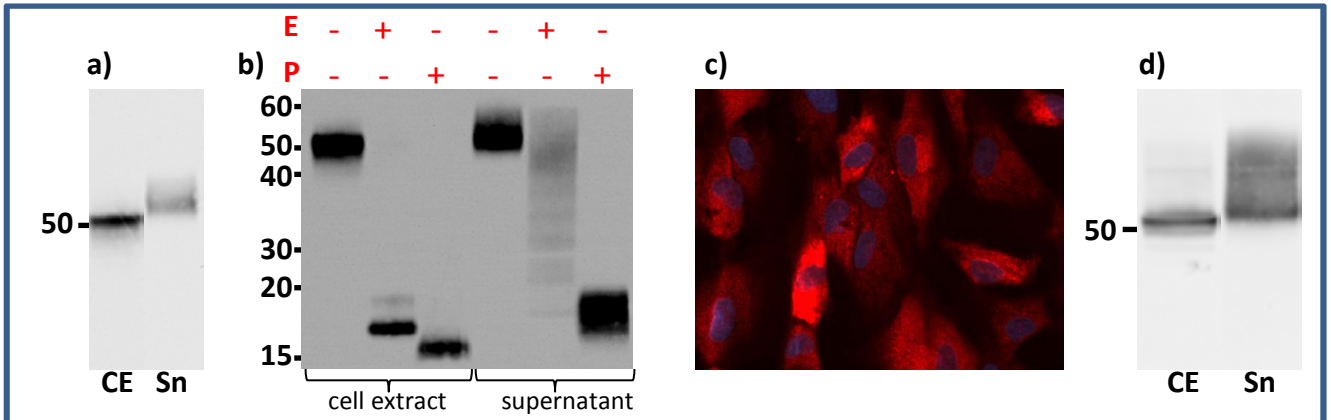


Figure 6.1: RAd-UL4 expression and glycosylation analysis

a) Western blot analysis of cell extracts (CE) and supernatants (Sn) of RAd-UL4-infected HFFF-hCARs (3 days p.i., MOI 10) using a V5 antibody. Loading in CE and Sn lanes represents 10%, and 0.3% of total samples respectively. Lanes were not adjacent lanes in the original blot and were pasted together. b) RAd-expressed gpUL4 deglycosylation with EndoH and PNGaseF enzymes analysed by Western blot using the monoclonal V5 antibody. Cell-associated and supernatant gpUL4 were mock-treated, or deglycosylated with EndoH (E) or PNGaseF (P) enzymes. c) Intracellular RAd-UL4 expression analysed by immunofluorescence as described in figure 3.3. Mock-infected controls are not shown. Nuclei were stained with DAPI (blue),whilst gpUL4 expression was detected using a V5 antibody (red). d) Western blot analysis of cell extracts (CE) and supernatants (Sn) of RAd-UL4-infected HFFF-hCARs (3 days p.i., MOI 10) using polyclonal sera. Loading in CE and Sn lanes represents 10%, and 0.3% of total samples respectively. Lanes were not adjacent lanes in the original blot and were pasted together.

6.3. gpUL4 in the context of HCMV infection

I sought to investigate whether gpUL4 is secreted in the context of HCMV infection. Antisera were generated against gpUL4 by immunisation of mice with RAd-UL4 (kindly performed by Dr Simone Cuff). This polyclonal antibody was able to detect gpUL4 in the cell extract or culture medium of cells infected with the Ad recombinant (Figure 6.1d), but not in strain Merlin-infected cell lysates or supernatants (data not shown). This indicated either low expression levels of gpUL4 in HCMV-infected cells or differential post-translational modifications during HCMV infection resulting in 'cloaking' of antigenic determinants.

Encouraged by success in detecting V5-epitope tagged proteins using the V5 antibody (see Chapters 3 and 5); I decided to generate a recombinant Merlin virus encoding V5-tagged UL4. A V5 epitope was fused to the C-terminus of the UL4 gene in the Merlin BAC as described in section 5.2 for genes UL74A and UL150A, and the recombinant virus designated RCMV1567. A time-course western blot analysis of gpUL4 expression was then performed using the V5 antibody on lysates and supernatants of HFFFs infected with RCMV1567, alongside HFFF-hCARs infected with RAd-UL4 (Figure 6.2). To remove contaminating virions, RCMV1567 supernatant samples were passed through 0.1 µm filters (Milliopore, SLVV033RS) prior to loading onto gel. Analysis of RCMV1567-infected cell extracts and supernatants detected gpUL4 in cell lysates at 1 day p.i. in the form of a sharp band migrating at 48 kDa; soluble gpUL4 could not be detected at this time-point. At 3 days p.i., gpUL4 was detected relatively abundantly in both cell lysates and supernatants. Supernatant gpUL4 migrated as a 55-60 kDa doublet, significantly more slowly than the cell-associated form. Interestingly, the difference in weight between the secreted and retained species was greater than that observed with RAd-UL4 infections. This demonstrates that secreted gpUL4 is more extensively modified than intracellular gpUL4, and that other viral element(s) are required for the completion of this action. By day 6, secreted and intracellular gpUL4 increased in abundance. A minor 60 kDa product was also expressed in cell lysates at day 6. The expression pattern of RAd-UL4 was as previously observed, except at 6 days p.i., several faster-migrating, and one 60 kDa slower-migrating species were also detected. The smaller proteins likely represent breakdown products; the larger band was expressed at very low levels.

Protein concentration was not normalised between cell extracts and supernatants in Figure 6.2a; cell extract samples were 20 times more concentrated by volume. Western blot was therefore also performed on samples containing equivalent proportions by volume, of cell

extract and supernatant preparations (Figure 6.2b). Similar levels of RCMV1567-expressed gpUL4 were present in cell lysates and supernatants at 3 days p.i.. However, at 6 days p.i., the bulk of gpUL4 appeared accumulated in the culture medium of cells infected with RCMV1567, indicating gpUL4 is primarily destined for secretion. gpUL4 expressed by RAd-infected cells also accumulated extracellularly. In these cells, intracellular gpUL4 was observed after longer exposure times (data not shown).

In order to characterise the glycan modifications of gpUL4 in the context of HCMV infection, deglycosyaltion analysis was performed on RCMV1567-infected cell lysates and supernatants at 6 days p.i. (Figure 6.3a). The major 48 kDa species was sensitive to EndoH whist treatment with this enzyme did not affect the minor 60 kDa form. Both species were sensitive to PNGaseF, migrating at near the predicted 17 kDa size of the primary translation product. As seen with RAd-UL4 (Figure 6.1b), secreted gpUL4 was partially sensitive to EndoH, migrating in a smear/ladder-like pattern; PNGaseF treatment produced a 17-21 kDa product.

Intracellular localisation of gpUL4 in the context of Merlin infection was also investigated (Figure 6.3b). Immunofluorescently detected gpUL4 aggregated in bright cytoplasmic foci near nuclei of infected cells at day 1 p.i.. At 3 days p.i., expression was more granular and less distinct, but still limited to perinuclear sites. Visibly stronger gpUL4 expression was observed at 6 days p.i..

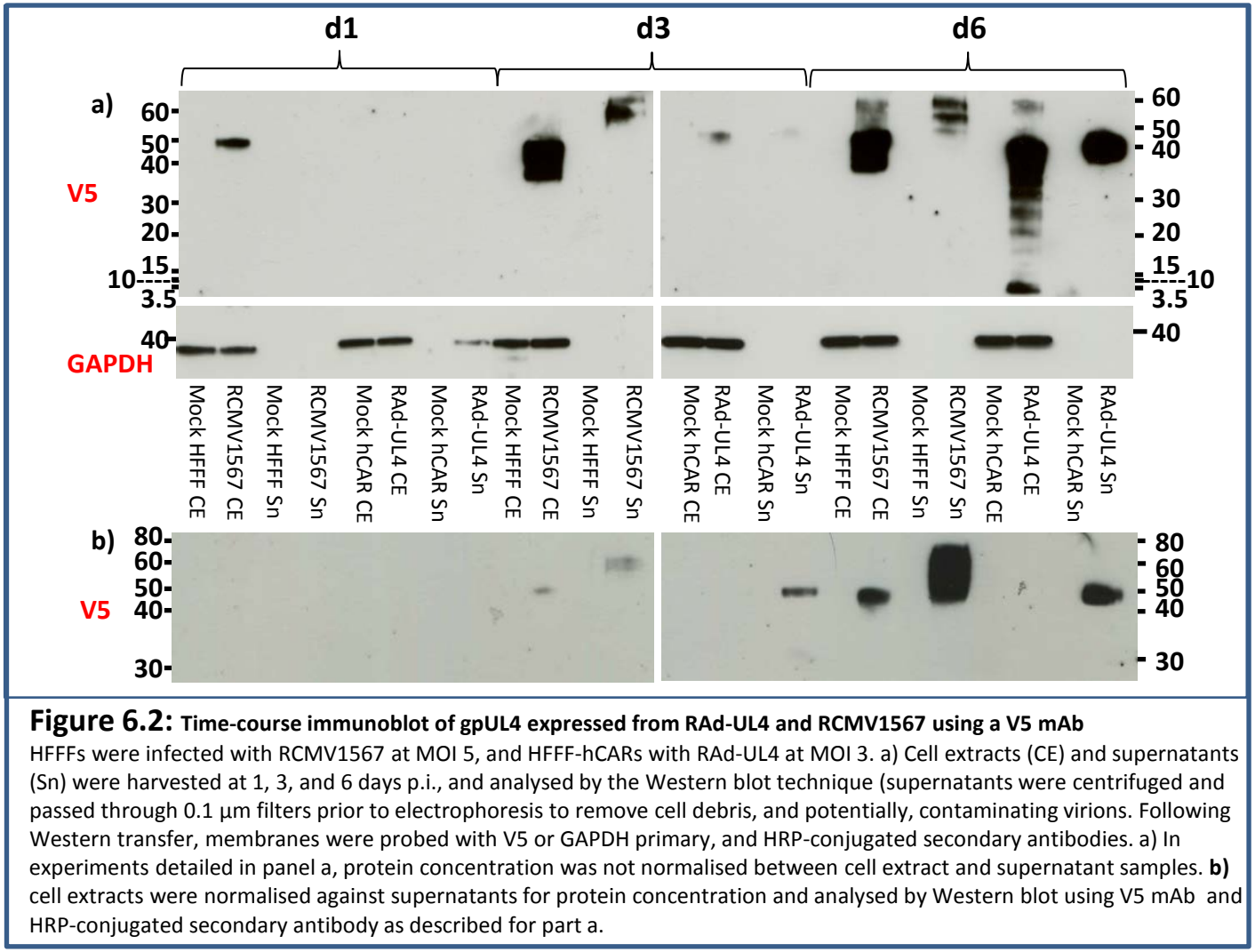


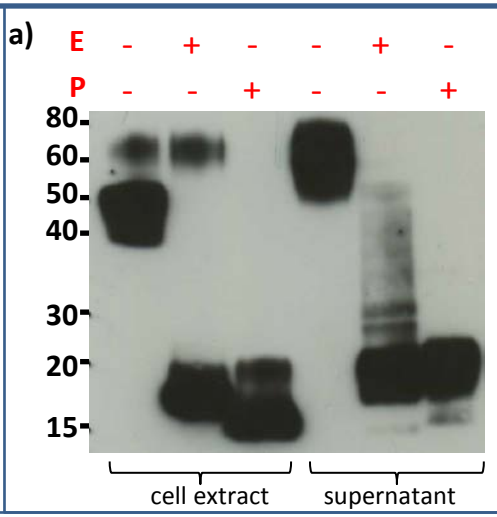
Figure 6.2: Time-course immunoblot of gpUL4 expressed from RAd-UL4 and RCMV1567 using a V5 mAb

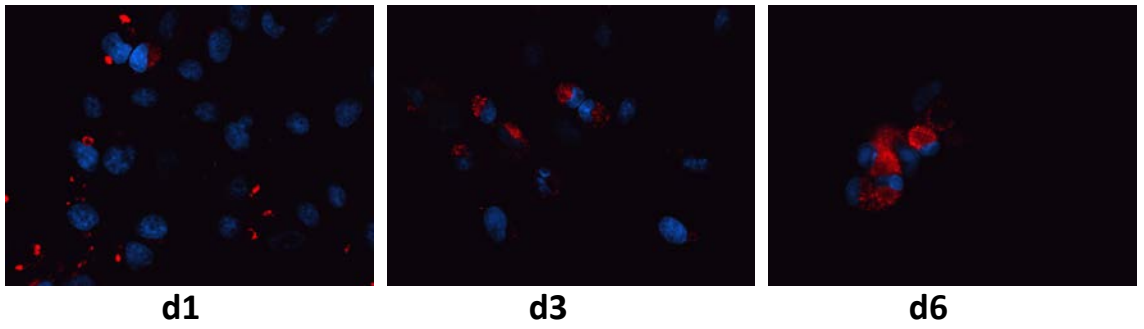
HFFFs were infected with RCMV1567 at MOI 5, and HFFF-hCARs with RAd-UL4 at MOI 3. a) Cell extracts (CE) and supernatants (Sn) were harvested at 1, 3, and 6 days p.i., and analysed by the Western blot technique (supernatants were centrifuged and passed through 0.1 µm filters prior to electrophoresis to remove cell debris, and potentially, contaminating virions. Following Western transfer, membranes were probed with V5 or GAPDH primary, and HRP-conjugated secondary antibodies. a) In experiments detailed in panel a, protein concentration was not normalised between cell extract and supernatant samples. **b)** cell extracts were normalised against supernatants for protein concentration and analysed by Western blot using V5 mAb and HRP-conjugated secondary antibody as described for part a.

Figure 6.3:

Glycosylation and intracellular localisation of gpUL4 in HCMV infected cells

a) HFFFs were infected with RCMV1567 at MOI 5. At 6 days p.i., cell-associated and supernatant gpUL4 were mock-treated, or deglycosylated with EndoH (E) or PNGaseF (P) enzymes. b) HFFFs were infected with RCMV1567 at MOI 5, fixed and permeabilised at indicated times, and gpUL4-V5 detected using a monoclonal V5 antibody. DAPI staining was performed to visualise infected-cell nuclei (blue).





6.4. gpUL4 is not concentrated in the virion

Chang et al. (1989) detected gpUL4 in purified extracellular viral particles and proposed that gp48 is a structural protein, although it was not specified whether gpUL4 was found in virions or dense bodies, as these particles were apparently not segregated in the purification process.

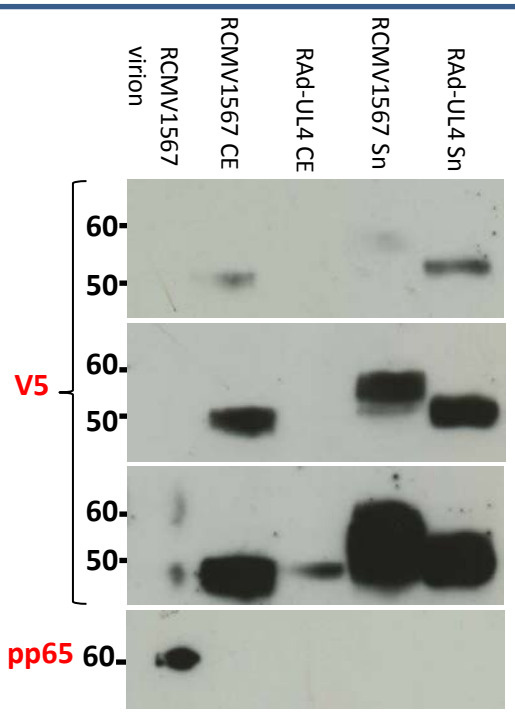
The Chang et al. study lacks important controls required to adduce specific and selective gpUL4 incorporation into virions. For example the inclusion of a cellular protein known not to be virion-associated would indicate the levels of contamination in the purified particle sample. However, the authors recognised that in contrast with other structural proteins, gpUL4 levels in virions and dense bodies were low, and that UL4 transcripts did not accumulate to high concentration at late stages of infection. Also, gpUL4 has not been confirmed as a virion protein in subsequent studies on the virion proteome. In a quantitative mass spectrometry analysis of protein species present in strain AD169 virions and dense bodies conducted by Varnum et al., gpUL4 levels did not meet the criteria of designation as a virion protein (Varnum et al., 2004).

I therefore aimed to determine whether gpUL4 is a virion component. RCMV1567 virions were purified on a NaT gradient, and analysed by the western blot technique in parallel with lysates and supernatants of cells infected with the said virus. Extracts and supernatants of cells infected with RAd-UL4 were also included. Following prolonged exposure to film, very weak gpUL4 expression was observed in purified virions compared to cell extract and supernatant samples (Figure 6.4). Conversely, pp65, a known tegument protein, was greatly concentrated in virions. These data suggest gpUL4 is not specifically packaged into virions; minute amounts of gpUL4 detected in purified virion preparations likely represent low-level contamination with intracellular or soluble gpUL4.

Figure 6.4:

gpUL4 is not concentrated in the virion

HFFFs were infected with RCMV1567 at MOI 5, and HFFF-hCARs with RAd-UL4 at MOI 3. Cell extracts (CE) and supernatants (Sn) were harvested at 3 days p.i., normalised by concentration, and analysed by immunoblotting alongside purified RCMV1567 virions. Prior to electrophoresis, supernatants were centrifuged and passed through 0.1 µm filters to remove cell debris and contaminating virions respectively. Following Western transfer membranes were probed with V5 or pp65 mAbs, and HRP-conjugated secondary antibody before exposure to film. 3 exposures of the same blot are shown for V5 mAb.



6.5. Soluble gpUL4 inhibits NK cell degranulation

The UL4 gene is non-essential for virus replication *in vitro* (Dunn et al., 2003b). Nevertheless, its abundant secretion by HCMV raises the possibility that it may be an immune-modulator, such as a virokine, chemokine/cytokine binding protein (CKBP), or NK-receptor-binding protein. With a view to performing immune-modulation assays, I proceeded to prepare concentrated soluble gpUL4. gpUL4 was purified from supernatants of cells infected with RAd-UL4 and RCMV1567 using a V5-tag purification kit as described in section 2.14.

To directly investigate the effect of gpUL4 on NK cell function, CD107a mobilisation assays were performed on HFFF-hTERTs in the presence or absence of purified soluble gpUL4 (Dr Rebecca Aicheler). Receptor-mediated recognition of target cells induces lytic granules within NK cells to migrate towards the site of interaction (Kuhn and Poenie 2002). The outer membrane of granules then merges with the plasma membrane, releasing its contents into the synaptic space in a process termed 'degranulation'. Degranulation results in exposure of CD107a, a glycoprotein normally present on lytic granule membranes, on the surface of NK cells. Surface-exposed CD107a thus serves as a marker of NK cell activation, and its detection using specific antibodies allows assessment of cell-mediated cytotoxicity (Betts et al. 2003).

As shown in Figure 6.5, 12.4% of NK cells degranulated in response to control HFFF-hTERTs, whilst 9.55% and 9.35% respectively degranulated when RAd-soluble gpUL4 and Merlin-soluble gpUL4 were included. This represents 23% and 24% inhibition of degranulation compared to control cells. The results of this assay indicate soluble gpUL4 is an NK-inhibiting molecule.

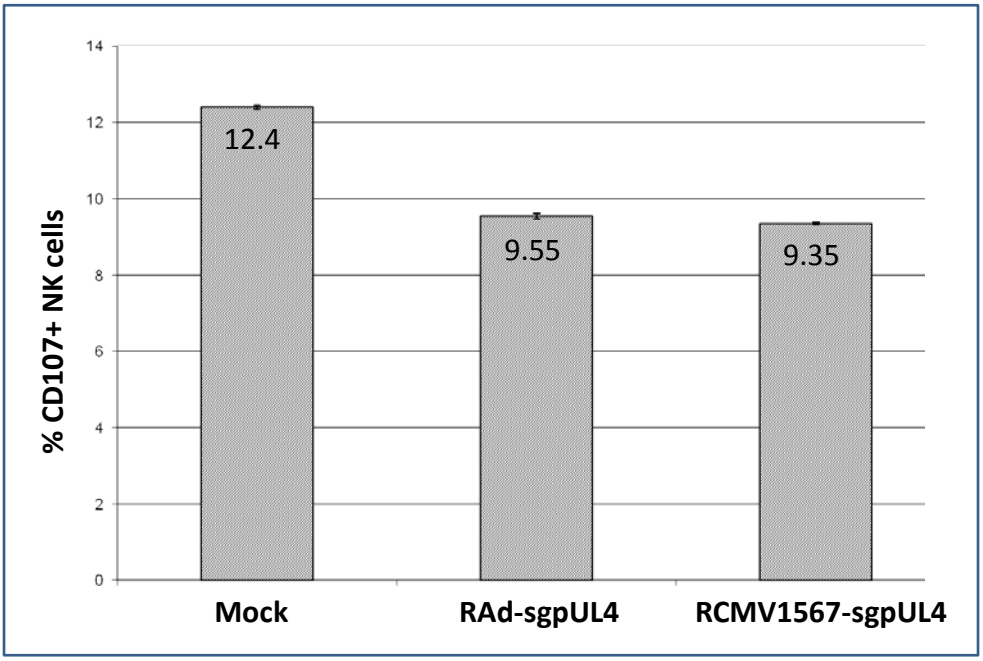


Figure 6.5:

Soluble gpUL4 inhibits NK cell degranulation in CD107a mobilisation assays

$5x10^5$ IFN-α-activated PBMC (D73) were incubated with $1x10^5$ HFFF-hTERTs and anti-CD107a-FITC mAB (BD Biosciences) for 6hr in either just media or media with purified sgpUL4 expressed from RAd-UL4 or Merlin-1567 (50 ng/ml final concentration). GolgiStop (BD Biosciences) was added for the last 5hr. After 6 hours, cells were stained for CD3 and CD56 and analysed on an Acurri flow cytometer. The percentages of CD107+, CD3-,CD56+ NK cells were calculated using Cflow software. Results are presented as means ± SEM of duplicate samples.

6.6. Summary

It was demonstrated that gpUL4 is abundantly secreted when expressed in isolation from an Ad recombinant, and in the context of HCMV Merlin infection. The observed kinetics of expression of intracellular gpUL4 are in agreement with the report by Chang et al (1989): gpUL4 is expressed at low levels during early stages of infection, accumulating to higher levels at late infection times. Compared with cell-associated gpUL4, the secreted isoform is of higher molecular weight and produced in greater abundance late in infection. Deglycosylation analysis revealed differences in the nature of carbohydrate moieties acquired by cell-associated and extracellular gpUL4. Moreover, western blot analysis of purified virions argues gpUL4 is not a virion component. CD107 mobilisation assays indicate soluble gpUL4 inhibits NK cell activation, and hence may play an important role in determining pathogenesis *in vivo*. Higher concentrations of soluble gpUL4 are currently being prepared to further investigate the immune-modulating potential of gpUL4.

7- Discussion

7.1. Cloning and expression of HCMV genes by Using an Adenovirus Vector

Ad vectors have been invaluable in the study of HCMV gene expression and function. Immuno-modulatory functions were discovered for the products of UL18, UL40, and UL142 when these genes were expressed using Ad vectors (Prod'homme et al., 2007, Tomasec et al., 2000, Wills et al., 2005a). Faithful to their functions during productive HCMV infection, RAd-derived gpUS2, gpUS3, gpUS6, and gpUS11 each independently downregulated cell surface MHC class-I expression (Hegde et al., 2002, Rehm et al., 2002, Tomasec et al., 2000, Tomazin et al., 1999). Also using this vector system, gpUS2 and gpUS3 were also able to inhibit MHC class-II antigen presentation, providing an escape route from recognition by $CD4^+$ T cells (Hegde et al., 2002, Tomazin et al., 1999). Huber et al. reported that low level expression of genes US7-US11 during productive HCMV infection hampered characterisation of their respective products; cloning of these genes into Ad vectors allowed expression to be detected (Huber et al., 2002). Co-expression of gH and gL from RAds was shown to enhance cell:cell fusion in HCMV-infected cells whilst co-infection of RAds expressing gH, gL and gO was shown to be necessary and sufficient for the formation of the gCIII complex (Milne et al., 1998, Paterson et al., 2002). Furthermore export of the gH/gL complex from the ER to the Golgi was enhanced in the presence of UL128, UL130, and UL131A when all five glycoproteins were expressed using RAds (Ryckman et al., 2008). Despite such examples, high-throughput generation of RAds for systematic screening of the HCMV genome was delayed by the technical limitations of standard vectors. Generation of the AdZ vector and the Merlin BAC facilitated large scale cloning of HCMV genes. This thesis describes the generation of 169 RAds each encoding a single HCMV gene, characterisation of expression of the cloned genes, and use of this RAd library in functional assays for the first time. Cloning of genes was carried out by the recombineering method; genes were PCR-amplified from the Merlin BAC and inserted into AdZ in a single step. Recombineering of 15 genes was extremely inefficient and required the insertion of a selectable marker (*galK*) prior to insertion into AdZ. The relatively GC-rich base composition of these genes compared to the rest of the HCMV genome correlated with their inefficient cloning.

Following the cloning of all the canonical HCMV genes into Ad vectors, it was important to evaluate their capacity to provide for transgene expression. Expression was validated for 160 out of 168 (95%) protein-coding genes using a combination of immunofluorescence (92% positive) and western blot (90% positive) analyses.

Expression of gB and gO was only detected after codon-optimised sequences of these genes were cloned into RAds. Optimisation of codon usage is known to facilitate high-level gene expression, and is being extended to the 8 genes whose expression was not detected. Of course this is not guaranteed to solve the problem; lack of detection may have been caused by other factors, such as cleavage of the C-terminal tag, or protein folding in a manner antagonistic to antibody binding. C-terminal cleavage is unlikely in the case of protein products of RL6, UL33, US9 and US14 however, because these proteins have been detected using C-terminal epitope tags (Salsman et al., 2008).

Excellent agreement was found between expression data presented here and those in comparable large scale expression analyses of HCMV genes; although the success rate in detecting transgene expression was higher with AdZ. Salsman et al. (2008) detected expression of 129 of 140 (87%) plasmid-cloned HCMV ORFs tested by immunofluorescence following transient transfection of 293T cells, whereas when HCMV ORFs were inserted into a baculovirus expression vector, western blot analysis revealed expression of only 50% of ORFs in infected insect cells (Gao et al., 2005).

7.1.1. Western blot analysis of expressed HCMV gene products

Predominant western blot species were detected within 10 kDa of predicted/reported sizes for >94% of constructs (128 of 136). Some interesting observations derived from this analysis are discussed:

Adenovirus replication takes place in the nucleus, so the capacity exists for primary transcripts to be spliced by endogenous host cell RNA processing systems; indeed the expression of intron-containing transcripts is known to be favoured. While recent studies indicated RNA splicing may be much more common than previously believed, it appears to be a relatively rare occurrence in the HCMV canonical ORFs. UL112 is known encode 34 kDa, 43 kDa, 50 kDa and 84 kDa phosphoproteins from a single alternatively spliced transcript (Wright and Spector, 1989). Products of almost identical size were detected in cells infected with RAd-UL112. Similarly, 22 kDa, 17 kDa, and 3.5 kDa proteins are synthesised from alternatively spliced US3 transcripts (Shin et al., 2006, Tenney et al., 1993). Consistent with this finding, 18 kDa and 16

kDa species were detected in RAd-US3 infected cells, as well as a minor species migrating closely to the 3.5 kDa molecular marker protein.

Western blot analysis identified several US12 family members as potential ubiquitin-conjugated proteins, due to the presence of high molecular-weight complexes characteristic of ubiquitinated proteins. This prediction was confirmed in the case of pUS20 (Dr Ceri Fielding, personal communication), and is being tested for US12, US13, US16, US18 and US21. Interestingly, western blot analysis of pUS16, pUS18 and pUS20 in the context of HCMV infection identified only protein singlets migrating near the predicted sizes of the primary translation products (Bronzini et al., 2012, Guo and Huang, 1993). This indicates ubiquitination of these proteins may be suppressed by other viral co-factors during productive infection, potentially to prevent their lysosomal/proteasomal targeting and degradation by the host cell.

HCMV is predicted to encode at least 50 glycoproteins (Table 14.1; Davison and Bhella 2007). Due to increase in mass resulting from glycosylation, the majority of these migrated significantly more slowly by SDS-PAGE than indicated by calculating the mass of the primary translation product. Of particular interest in the study, 11 out of 16 predicted, yet previously uncharacterised glycoproteins were shown to encode products larger than the predicted primary translation product. Deglycosylation analysis provided confirmation that RL12, UL6 and UL7 were glycoproteins. The remaining 5/16 predicted glycoproteins migrated at or below the predicted molecular mass of their primary translation products; this does not rule out glycosylation, since these products may be subject to proteolytic cleavage. There would be merit in undertaking systematic analyses of HCMV gene products expressed using the Ad vector for posttranslational processing.

7.1.2. Intracellular localisation of expressed HCMV gene products

The intracellular localisation of 98 out of 109 gene products (90%) expressed by using the AdZ vector was consistent with previous descriptions. In the case of the other 11 genes, products of 9 had been described solely in the context of productive HCMV infection. These proteins may well require interaction with additional HCMV-encoded proteins for appropriate trafficking. The fact that 8 of the 9 proteins in this category are nuclear during productive HCMV infection but cytoplasmic using Ad recombinants indicates a large proportion of nuclear HCMV-encoded proteins may lack their own nuclear localisation signals (NLS). However, it cannot be discounted that for some, nuclear localisation could have been perturbed by the C-terminal V5 tag. While analysis of the RAd library showed UL89 and UL104 to be cytoplasmic, during lytic infection, pUL89 and pUL104 both interact and co-localise with pUL56 in the

nucleus (Thoma et al., 2006). pUL56 in turn is nuclear-localised in the context of HCMV infection and during isolated expression; nuclear translocation of pUL89 and pUL104 may therefore require co-expression with pUL56. Interestingly, a reversal of this possible scenario occurs in HSV-1, where during independent expression, the pUL56 homologue (pUL28) is cytoplasmic whilst pUL89 and pUL104 homologues (UL15 and UL6 respectively) are nuclear. Co-expression with either HSV-1-UL15 or HSV-1-UL6 was sufficient to translocate HSV-1 pUL28 to the nucleus (Abbotts et al., 2000, Koslowski et al., 1999, White et al., 2003). The prediction that pUL56 mediates nuclear translocation of pUL89 and pUL104, needs to be tested experimentally.

The other 2 inconsistencies with regards protein localisation involve pUL80.5 and gpUL130. RAd-expressed pUL80.5 staining was punctate PML body-like within the nucleus but has been reported as pan-nuclear in Cos-7 cells (Plafker and Gibson, 1998, Wood et al., 1997), however, Cos-7 cells are both derived from African Green Monkey Kidneys (i.e. not human) and have been transformed with the SV40 Large T antigen. gpUL130 expression was previously visualised throughout the cytoplasm and on the surface of transfected 293T cells (Gerna et al., 2008a). RAd-expressed gpUL130 however, was present only in the nucleus-proximal portion of HFFF-hCARs, clearly excluded from the cell membrane. The pentomeric gH/gL/UL128/UL130/UL131A complex is packaged into the virions of low passage isolates to promote endothelial cell tropism. gL has long been known to provide a chaperon function that is required for the appropriate folding of gH, and its subsequent trafficking to the plasma membrane and ultimately into virus particles. gH is the only component of the pentomeric complex that has a transmembrane domain and thus plays a key role in anchoring the complex in membranes. The pentomeric complex provides a clear example of the need for HCMV genes to be co-expressed in order to reconstitute multiprotein complexes.

7.2. Regulation of Fas expression by HCMV

HCMV infection was demonstrated to result in efficient downregulation of cell surface Fas expression that is detectable at early times p.i. (24 hrs), and persists through the late phase. The reduction in Fas expression was not mediated by a secreted factor, but was dependent on *de novo* virus-encoded gene expression.

Fas downregulation was deployed as a pilot study to establish a strategy for mapping HCMV gene function. As a first step, a series of 10 HCMV deletion mutants each lacking a block of non-essential genes was tested in a loss-of-function assay. Similar levels of Fas downregulation were observed between all deletion mutants and wildtype HCMV Merlin, suggesting none of the 58 genes absent in these mutants is essential for Fas downregulation by strain Merlin. The limitations of this approach include: (i) only approximately 35% of the HCMV canonical genes are deleted, (ii) some of the mutants are replication impaired; this can alter the kinetics of specific virus-encoded effects, and (iii) a redundant function may give only a partial or no phenotype in the screen.

Subsequently, gain-of-function screening of the panel of Ad recombinants encoding HCMV genes also did not identify the gene(s) responsible for Fas downregulation. Having characterised the expression of some 170 Ad recombinants, there are a number of reasons why this strategy also could miss a prescribed function: (i) the gene responsible may not have been annotated and thus not be present in the panel of Ad recombinants. (ii) The function may be encoded by a miRNA; the initial strategy for generating the RAd library also included insertion of HCMV miRNAs; indeed an AdZ vector has been utilised in RNA interference (McLaren et al., 2010). However, this component was not prioritised as there was no immediate prospect of validating RAd-miRNA constructs. The AdZ vector technology is currently being optimised for this application. (iii) More than one HCMV gene may be required for the function. While systematically analysing the expression of Individual functions for Table 3.2, it became apparent that a number of HCMV genes act in concert in a reliable and reproducible fashion (e.g. the members of the UL128L pentomeric complex). An advantage of the Ad vector system is that it permits multiple constructs to be co-delivered. There is

therefore merit in developing subroutines in the gene screening strategy that allow sets of co-operating viral functions to be tested together. With this objective in view, it would be worth investigating which pathways could be recapitulated using Ad vectors e.g. glycoprotein complexes, DNA replication, capsid/dense body assembly. (iv) The Ad-cloned HCMV gene may be expressed at levels inappropriate for the function. The high level of gene expression that is usually achieved using Ad vectors may be antagonistic to function. Conversely, expression levels of some genes may be insufficient for function of interest. (v) Protein expression in an inappropriate compartment. Products of 13 genes are known to localise differently if expressed independently of other viral proteins: RL13, UL48A, UL53, UL70, UL73, UL74, UL75, UL82, UL86, UL94, UL100, UL114, and TRS1 (see Table 3.2). Moreover, viral factors that promote appropriate localisation of 9 of these gene products have been identified, but remain unknown in the cases of gpRL13, pUL70, pp71 (UL82), and TRS1. As with genes that act cooperatively to encode a function, co-delivery of Ad recombinants can be utilised to promote appropriate intracellular localisation. (vii) The C-terminal V5 tag may disable the function.

Comparable levels of Fas downregulation were observed on the surface of cells infected with strains Merlin, AD169, FIX, or TB40; this assay was performed using fibroblast cells of various origins. These results demonstrated that Fas downregulation is a conserved function between different HCMV variants, and traits specific to a particular batch of cells were not responsible for initial observation of this function. In this respect, an independent confirmation of Fas downregulation during HCMV infection was provided by SILAC analysis of cell surface molecules of fibroblasts (Dr Peter Tomasec personal communication).

Downregulation of Fas by different HCMV strains is indicative of a potentially important role for this novel function in viral replication and pathogenesis. With regards mapping the gene responsible for Fas downregulation, results of this assay allowed me to conclude that the 25 HCMV genes deleted/inactive in strains AD169, FIX and TB40 are not essential for Fas downregulation, effectively extending the number of discounted genes to 83.

Lytic infection of HFFs with strain AD169 was reported to cause a reduction in Fas mRNA levels (4.3 fold) (Browne et al., 2001). However, levels of Fas mRNA were not significantly altered during lytic infection with strain Merlin (data not shown); downregulation of Fas at the cell surface was therefore anticipated to occur at a posttranscriptional level. Western blot analysis showed that the clear decrease in Fas levels at the surface of infected cells was not accompanied by a global reduction in total Fas protein levels. Although modest fluctuations in total Fas expression were seen on days 2 and 3 p.i. in Figure 4.7a, overall, there was no clear evidence of the virus either stimulating or suppressing total Fas levels, even at 6 days p.i. (data

not shown). HCMV infection thus results in FAS being selectively downregulated at the cell surface, and retained in an intracellular compartment. This phenomenon could take place through either removal of existing Fas molecules from the cell surface, or blockage of transport of newly-synthesised Fas to the cell surface. Fas extracted from mock-infected or HCMV-infected cells was resistant to EndoH treatment, indicating ER sequestration is not the mechanism of retention, since proteins only acquire resistance to EndoH treatment once trafficked beyond the ER and into the Golgi apparatus.

Fas joins a growing list of cellular proteins that appear redistributed, rather than degraded during HCMV infection. These include ULBP2, MICB, CD155, TNFRI and TRAILR2 (Dunn et al., 2003, Stern-Ginossar et al., 2007b, Tomasec et al., 2005, Baillie et al., 2003, unpublished data). While signalling through Fas, TNFRI and TRAILR2 are primarily associated with the induction of apoptosis; in their unstimulated form they may provide pro-survival stimuli. Intracellular levels of TNFR1, another member of the TNF superfamily that is downregulated from the surface of HCMV-infected cells, are also not depleted during infection (Baillie et al., 2003). TNFR1 proteins were instead relocalised from the cell surface to the TGN, appearing in intensely-stained perinuclear regions. In addition, the close functional and familial connections between Fas and TNFRI dictate speculation that both proteins may be downregulated by the same mechanism, and/or the same HCMV gene(s).

Fas-signalling was stimulated in mock-infected and strain Merlin-infected cells using Fas ligand or a cross-linking Fas antibody; infected cells displayed significantly lower caspase-3/7 activity, signifying protection from apoptosis. This recapitulated previous findings that infection with HCMV renders cells more resistant to Fas-mediated apoptosis. Nevertheless, the contribution of Fas downregulation towards enhanced cell survival in this assay cannot be determined because two other HCMV genes, namely UL36 and UL37, also protect infected cells from Fas-mediated apoptosis. In this context, pUL36 inhibits activation of caspase-8 whilst pUL37 inhibits permeabilisation of the outer mitochondrial membrane by sequestering Bax within mitochindria, thus protecting infected cells from apoptosis induced by a range of different stimuli (Arnoult et al., 2004, Skaletskaya et al., 2001). Previously, it was assumed that HCMV-infected cells acquired resistance against Fas-mediated apoptosis first through pUL36 activity (24 hrs p.i.), and then through the activity of pUL37 (48 hrs p.i.). However, Fas downregulation is also complete at these time-points, and may contribute significantly to protection of infected cells from Fas-mediated apoptosis. A recombinant ΔUL36 Merlin virus was generated for the purpose of assessing the relative contributions of Fas downregulation and UL36 activity on Fas-mediated apoptosis by comparison with wildtype strain Merlin. However, this recombinant virus showed severe growth defect and could not be used experimentally.

Further investigation is required to determine the HCMV genes responsible for Fas downregulation, and the mechanism of this function.

Downregulation of Fas from the cell surface is not a trait specific to HCMV. The adenovirus receptor internalisation and degradation (RID) protein, also called the E3-10.4K/14.5K complex, inhibits apoptosis by promoting loss of Fas molecules from the surface of infected cells (Elsing and Burgert, 1998). Similar to observations made in the current study with HCMV, RID does not inhibit Fas expression at the transcriptional level but rather mediates internalisation, and subsequent degradation of existing Fas molecules by the lysosomal pathway (Tollefson et al., 1998). Contradictory reports have been published with regards total levels of Fas following internalisation and degradation by RID. In one study, intracellular levels of Fas were not significantly altered after infection with a recombinant adenovirus (Shisler et al., 1997), whilst another study demonstrated substantial decrease in levels of Fas in HCMV-infected cells (Tollefson et al., 1998). The ER-resident Myxoma virus leukemia-associated protein (MV LAP) downregulates Fas and MHC class-I molecules from the surface of transfected cells (Guerin et al., 2002). Interestingly, Fas levels at the surface of MV-infected cells were not affected, implying LAP acts to counter Fas upregulation that may be caused by another viral/viral-induced factor. Also of note, is that retention of MV LAP in the ER was crucial to its function in preventing Fas expression at the cell surface, yet downregulation of Fas or MHC class-I did not occur through ER retention. Abduction of cell surface Fas and MHC class-I by MV LAP likely involves exacerbation of endocytosis and subsequent lysosomal degradation, and has been proposed to occur through interactions with adaptor proteins of the actin cytoskeleton. Nevertheless, precise mechanisms underlying this function have yet to be elucidated. Identification of LAP homologues in other poxviruses and several gammaherpesviuses including HHV-8 has prompted the designation of these molecules as surface cellular receptor abductor proteins (scrapins). No homology was reported between scrapin family members and any HCMV genes. From an evolutionary viewpoint, conservation or commonality of function within or between species generally infers importance of function. Downregulation of Fas at the surface of virus-infected cells could potentially inhibit host-induced apoptosis, providing infected cells, and effectively the virus, with an escape route from recognition and elimination by the host's immune response.

7.3. Characterisation of the novel HCMV UL150A gene

Deep sequencing of the Merlin transcriptome was performed by Gatherer et al. (2011) to provide a better understanding of HCMV gene content and usage; the analysis was performed at 72 hours p.i. in fibroblasts. Great variation exists in the level of transcription across the genome, yet nearly every part of the genome is transcribed to some extent. Anti-sense transcripts were promiscuously produced from a large proportion of the genome, US33A was designated as a novel HCMV gene, and UL30A expression was authenticated. This study also uncovered the remarkable statistic that over two thirds of polyA-containing viral transcripts represent NNTs. Splicing was found to be much more common than had been appreciated, and the majority of previously recognised splice sites were confirmed. Most notable amongst the novel spliced transcripts, was the presence of 5 super acceptor sites (SASs), each one spliced from multiple upstream donors, and positioned upstream of a potential coding region. I investigated the expression of two novel SAS-associated coding regions, namely UL74A and UL150A, in the context of HCMV infection. V5 tags were fused to the C-termini of UL74A and UL150A in the Merlin BAC, and recombinant viruses generated (designated RCMV1569 and RCMV1571 respectively). Western blot and immunofluorescence analyses confirmed pUL150A was expressed and the production of multiple species was consistent with fusion to multiple upstream ORFs. pUL150A was produced with early/late kinetics, and localised to the ER. By western blot, pUL150A migrated as a major 34 kDa doublet, consistent with the predicted 31 kDa size of the primary translation product, although a minor 50 kDa isoform was also visible at late times of infection. Interestingly, the 34 kDa doublet became more abundant from 3 days p.i. to 6 days p.i., but the expression of the 50 kDa isoform remained constant, suggesting potentially distinct regulation of expression. The 50 kDa migrating species may represent an alternative splice variant with an extended 5' terminus, or posttranslationally processed pUL150A. Characterisation of UL150A expression marks the first example of a functional ORF that is positioned entirely antisense to another larger, unspliced coding region (UL150).

Strong evidence that UL74A is expressed as part of an alternatively spliced transcript is provided by northern blot, RT-PCR, and RACE data (Scalzo et al., 2009, Gatherer et al., 2011). However, when UL74A open reading frame was tagged its expression could not be detected by

either immunofluorescence or western blot. With tagged proteins, the possibility exists that the tag may be lost due to protein cleavage, or its antigenicity destroyed or concealed by folding of the fused protein. While tagging of HCMV ORFs is an excellent device for demonstrating and characterising expression, a negative result does not necessarily mean the gene is not expressed. Insertion of an epitope tag in a region other than the C-terminus of the predicted protein could enable detection.

7.4. Characterisation of gpUL4

gpUL4 was readily detected in the culture media of cells infected with RAd-UL4. Three lines of evidence suggested supernatant RAd-gpUL4 was not analogous to intracellular gpUL4, and was thus actively secreted. First, supernatant gpUL4 consistently migrated more slowly than the intracellular form in western blot, indicative of more extensive glycosylation. Second, supernatant gpUL4 was more resistant to EndoH treatment than the intracellular form, suggesting passage through the Golgi. Third, supernatant but not cell-associated gpUL4 showed some degree of resistance to PNGaseF treatment, typical for proteins with O-linked glycosylation.

Polyclonal sera prepared by immunising mice with RAd-UL4 detected RAd-expressed, but not Merlin-expressed gpUL4. I suspected this could have been caused by lower expression levels of gpUL4 in the context of HCMV infection. The strong MIE promoter in AdZ often drives far greater HCMV gene expression compared to endogenous levels. It is also possible that gpUL4 produced during HCMV infection undergoes distinct post-translational modifications and is not efficiently bound by the polyclonal sera.

As an alternative approach to characterise expression in the context of HCMV infection, UL4 was V5-tagged within the Merlin BAC. Using the V5 antibody, tagged gpUL4 was readily detected in lysates and supernatants of cells infected with the recombinant virus, confirming gpUL4 is a secreted HCMV protein. The migration and temporal expression pattern of intracellular gpUL4 were consistent with previous findings (Chang et al., 1989). Soluble gpUL4 appeared in the extracellular media only at late times of infection. Difference in time of expression between the two isoforms suggests complex temporal regulation of the UL4 promoter. Alternatively, other viral factors may inhibit secretion at earlier time-points. Soluble gpUL4 also migrated more diffusely, and considerably more slowly than cell-associated gpUL4. Interestingly, the size difference between intracellular and extracellular gpUL4 was greater than that observed using RAd-UL4. This adds gpUL4 to the list of HCMV glycoproteins which require other viral factors to undergo complete glycosylation: RL13, UL18, UL73, UL74 (Griffin et al., 2005, Mach et al., 2000, Stanton et al., 2010, Theiler and Compton, 2002).

In a previous investigation into gpUL4, inhibition of N-linked glycosylation did not lead to detection of the 17 kDa primary polypeptide from cells infected with strain Towne (Chang et

al., 1989). However, but for a high molecular-weight species expressed at true-late times, Merlin-derived intracellular gpUL4 was almost fully sensitive to EndoH treatment. It is possible that inter-strain differences between Merlin and Towne are responsible for variation in glyco-processing of gpUL4. While Chang et al. (1989) detected gpUL4 in purified virus particles; it was present in lower levels than other structural components. In a subsequent study by Varnum et al. (2004), levels of gpUL4 in the virion did not meet the qualifying threshold to be designated a structural protein. Here, it was shown that only minute amounts of gpUL4 were detected in purified Merlin virions compared to cell lysates and supernatants, suggesting low-level contamination rather than selective incorporation. Having now been defined as a secreted glycoprotein, it is natural to consider the possibility gpUL4 may be acting as a virokine, akin to cmvIL-10 or the IL-8-like function pUL146.

Merlin- and RAd-UL4-derived soluble gpUL4 purified on V5 antibody columns proved capable of inhibiting NK cell degranulation in CD107 mobilisation assays, suggesting a possible function as a virokine or chemokine/cytokine-binding protein receptor antagonist. The NK cell inhibiting function of soluble gpUL4 will be further characterised upon preparation of more concentrated samples. Planned experiments include screening a panel of leukocyte receptors to identify potential binding target(s). NK cell inhibitory function is likely specific to the secreted form of gpUL4 since RAd-UL4-infected cells did not inhibit NK cell activation relative to mock-infected cells (Dr James Davies, personal communication).

A BLAST search of the UL4 amino acid sequence revealed homology with CD180/RP105 (amino acids 51-121, 30% homology, 46% similarity). CD180 is a Toll-like receptor (TLR) homologue, and an important negative regulator of TLR4 signalling (Divanovic et al., 2005). TLR4 forms a complex with MD2 at the cell surface, and binds lipopolysaccharide (LPS), a constituent of the outer membrane of Gram-negative bacteria. Other TLR4 targets include viral components such as respiratory syncytial virus (RSV) fusion proteins, vesicular stomatitis virus (VSV) glycoprotein G, and Ebola virus (EV) glycoprotein (Georgel et al., 2007, Kurt-Jones et al., 2000, Okumura et al., 2010). HCMV is also known to activate signalling through TLR4, although the precise viral product(s) that are recognised by TLR4 have not been identified. It has been demonstrated that infection of monocytes with strain Towne induces TLR4 signalling (Yew et al., 2012). Stimulation of the TLR4 signalling pathway with LPS led to production of inflammatory cytokines, and inhibited HCMV infection in foreskin fibroblasts and ectocervical tissue (Sailesh et al., 2007). From the above findings, it can be seen that countering TLR4 signalling may promote HCMV survival within host cells.

Preliminary experiments were performed to determine whether soluble gpUL4 is able to interfere with TLR4 signalling. IL-1β production was measured by enzyme-linked immunosorbent assay (ELISA) from THP-1 monocytes incubated with LPS or LPS and soluble gpUL4, and compared to that from untreated controls. Merlin- and RAd-UL4-derived soluble gpUL4 significantly inhibited IL-1β production in the presence of LPS (data not shown), suggesting function as an inhibitor of TLR4 signalling. Although promising, description of these results is raised tentatively and must be viewed with caution, since the amounts of soluble gpUL4 initially purified were insufficient for repeat of this assay. More concentrated soluble gpUL4 is currently being prepared in order to further investigate the effect of gpUL4 on TLR4 signalling.

UL4 is amongst the class of genes that contain regulatory upstream open reading frames (uORFs). uORFs are *cis*-acting elements within the 5' leader region of some primary transcripts that play important roles in regulating translation of their respective main ORFs (Wethmar et al., 2010). uORFS are highly conserved between homologous genes; yet different uORFs are extremely variable in length, sequence, structural features and regulatory function. Bioinformatic analysis has shown that uORFs are present in diverse biological organisms such as fungi, viruses and mammalian cells. Although approximately half of human transcripts contain uORFs, this proportion is lower than might occur by random chance; furthermore uORF-mediated translational control has only been demonstrated for about 100 eukaryotic genes. Amongst these, there is a remarkable over-representation of key groups of genes encoding oncogenes, growth factors and cellular receptors (Morris and Geballe, 2000). The functional significance of uORFs in humans is reflected in the correlation between disrupting uORF mutations, and increased susceptibility to severe disorders such as thrombocythemia, cancer and Alzheimer's disease (Medenbach et al., 2011). Viral transcripts other than UL4 that are known to contain functional uORFs include the human immunodeficiency virus (HIV) type 1 mRNAs, hepatitis B virus (HBV) pregenomic RNA, and tomato golden mosaic virus AL2 and AL3 mRNAs (Chen et al., 2005, Luukkonen et al., 1995, Shung and Sunter, 2009).

Chang et al. (1989) first described the presence of 3 uORFs in the 5' leader region of strain Towne UL4 primary transcript. Although this 5' leader region inhibited translation of the major UL4 ORF in cell free assays, deletion analysis suggested none of the 3 uORFs was essential for this inhibition. In a series of elegant studies using intact cells and cell-free assays, it was subsequently demonstrated that strain Towne uORF-2 (i) is essential for inhibition of gpUL4 translation, (ii) is translated into a polypeptide product that remains ribosome-associated, and (iii) causes translation inhibition by sequence-dependent ribosomal stalling (Schleiss et al., 1991, Degnin et al., 1993, Cao and Geballe, 1995, Cao and Geballe, 1996, Alderete et al., 1999).

It was also shown that deletion or mutagenesis of some but not all of the 22 uORF-2 amino acids de-repressed gpUL4 translation. The uORF-2 sequence is well conserved amongst HCMV species; strains encoding non-functional uORF2 consistently express higher levels of gpUL4. Merlin uORF2 differs from that of Towne by a single T16I mutation. This particular mutation was noted in 5 other clinical isolates and is not thought to alter uORF2 function (Alderete et al., 1999).

The high-degree of inter-strain conservation in the amino acid sequence of gpUL4 and its associated uORF-2; and potential multi-dimensional properties of soluble gpUL4 in evading the immune response point towards the importance of this HCMV gene product *in vivo*.

8- References

ABBOTTS, A. P., PRESTON, V. G., HUGHES, M., PATEL, A. H. & STOW, N. D. 2000. Interaction of the herpes simplex virus type 1 packaging protein UL15 with full-length and deleted forms of the UL28 protein. *Journal of General Virology,* 81**,** 2999-3009.

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. *Journal of General Virology,* 83**,** 1315-1324.

ADAIR, R., LIEBISCH, G. W. & COLBERG-POLEY, A. M. 2003. Complex alternative processing of human cytomegalovirus UL37 pre-mRNA. *Journal of general virology,* 84**,** 3353-3358.

ADAM, B. L., JERVEY, T. Y., KOHLER, C. P., WRIGHT, G., NELSON, J. A. & STENBERG, R. M. 1995. The human cytomegalovirus UL98 gene transcription unit overlaps with the pp28 true late gene (UL99) and encodes a 58-kilodalton early protein. *Journal of virology,* 69**,** 5304-5310.

ADAMO, J. E., SCHRÖER, J. & SHENK, T. 2004. Human cytomegalovirus TRS1 protein is required for efficient assembly of DNA-containing capsids. *Journal of virology,* 78**,** 10221-10229.

ADLER, B., SCRIVANO, L., RUZCICS, Z., RUPP, B., SINZGER, C. & KOSZINOWSKI, U. 2006. Role of human cytomegalovirus UL131A in cell type-specific virus entry and release. *Journal of General Virology,* 87**,** 2451-2460.

ADLER, S. 1990. Cytomegalovirus and Child Day Care: Evidence for An Increased Infection Rate Among Day Care Workers. *The Pediatric Infectious Disease Journal,* 9**,** 862.

ADLER, S. P. 1989. Cytomegalovirus and child day care. *New England Journal of Medicine,* 321**,** 1290-1296.

ADLISH, J. D., LAHIJANI, R. S. & JEOR, S. C. S. 1990. Identification of a putative cell receptor for human cytomegalovirus. *Virology,* 176**,** 337-345.

AGGARWAL, B. B., SINGH, S., LAPUSHIN, R. & TOTPAL, K. 1995. Fas antigen signals proliferation of normal human diploid fibroblast and its mechanism is different from tumor necrosis factor receptor. *FEBS letters,* 364**,** 5-8.

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.

AHN, K., ANGULO, A., GHAZAL, P., PETERSON, P. A., YANG, Y. & FRÜH, K. 1996. Human cytomegalovirus inhibits antigen presentation by a sequential multistep process. *Proceedings of the National Academy of Sciences,* 93**,** 10990.

AHN, K., GRUHLER, A., GALOCHA, B., JONES, T. R., WIERTZ, E. J. H. J., PLOEGH, H. L., PETERSON, P. A., YANG, Y. & FRÜH, K. 1997. The ER-luminal domain of the HCMV glycoprotein US6 inhibits peptide translocation by TAP. *Immunity,* 6**,** 613-621.

AKRIGG, A., WILKINSON, G. & ORAM, J. 1985. The structure of the major immediate early gene of human cytomegalovirus strain AD169. *Virus Research,* 2**,** 107-121.

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. G. 2003. Two novel spliced genes in human cytomegalovirus. *Journal of General Virology,* 84**,** 1117.

AL-BARAZI, H. O. & COLBERG-POLEY, A. M. 1996. The human cytomegalovirus UL37 immediate-early regulatory protein is an integral membrane N-glycoprotein which traffics through the endoplasmic reticulum and Golgi apparatus. *Journal of virology,* 70**,** 7198-208.

ALDERETE, J. P., JARRAHIAN, S. & GEBALLE, A. P. 1999. Translational effects of mutations and polymorphisms in a repressive upstream open reading frame of the human cytomegalovirus UL4 gene. *Journal of virology,* 73**,** 8330-8337.

ALFORD, C. A., STAGNO, S., PASS, R. F. & BRITT, W. J. 1990. Congenital and perinatal cytomegalovirus infections. *Review of Infectious Diseases,* 12**,** S745.

ALVISI, G., RIPALTI, A., NGANKEU, A., GIANNANDREA, M., CARAFFI, S. G., DIAS, M. M. & JANS, D. A. 2006. Human cytomegalovirus DNA polymerase catalytic subunit pUL54 possesses independently acting nuclear localization and ppUL44 binding motifs. *Traffic,* 7**,** 1322-1332.

ANDERS, D., IRMIERE, A. & GIBSON, W. 1986. Identification and characterization of a major early cytomegalovirus DNA-binding protein. *Journal of virology,* 58**,** 253-262.

ANDERS, D. G. & GIBSON, W. 1988. Location, transcript analysis, and partial nucleotide sequence of the cytomegalovirus gene encoding an early DNA-binding protein with similarities to ICP8 of herpes simplex virus type 1. *Journal of virology,* 62**,** 1364-1372.

ANDREWES, C., BURNET, F., ENDERS, J., GARD, S., HIRST, G., KAPLAN, M. & ZHDANOV, V. 1961. Taxonomy of viruses infecting vertebrates: present knowledge and ignorance. *Virology,* 15**,** 52.

ARLT, H., LANG, D., GEBERT, S. & STAMMINGER, T. 1994. Identification of binding sites for the 86-kilodalton IE2 protein of human cytomegalovirus within an IE2-responsive viral early promoter. *Journal of virology,* 68**,** 4117-4125.

ARNON, T. I., ACHDOUT, H., LEVI, O., MARKEL, G., SALEH, N., KATZ, G., GAZIT, R., GONEN-GROSS, T., HANNA, J. & NAHARI, E. 2005. Inhibition of the NKp30 activating receptor by pp65 of human cytomegalovirus. *Nature immunology,* 6**,** 515-523.

ARNOULT, D., BARTLE, L. M., SKALETSKAYA, A., PONCET, D., ZAMZAMI, N., PARK, P. U., SHARPE, J., YOULE, R. J. & GOLDMACHER, V. S. 2004. Cytomegalovirus cell death suppressor vMIA blocks Bax-but not Bak-mediated apoptosis by binding and sequestering Bax at mitochondria. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 7988-7993.

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. *Journal of virology,* 83**,** 12345-12354.

ATALAY, R., ZIMMERMANN, A., WAGNER, M., BORST, E., BENZ, C., MESSERLE, M. & HENGEL, H. 2002. Identification and expression of human cytomegalovirus transcription units coding for two distinct Fcγ receptor homologs. *Journal of virology,* 76**,** 8596-8608.

AUCOIN, D. P., SMITH, G. B., MEIERING, C. D. & MOCARSKI, E. S. 2006. Betaherpesvirus-conserved cytomegalovirus tegument protein ppUL32 (pp150) controls cytoplasmic events during virion maturation. *Journal of virology,* 80**,** 8199-8210.

AWASTHI, S., ISLER, J. A. & ALWINE, J. C. 2004. Analysis of splice variants of the immediate-early 1 region of human cytomegalovirus. *Journal of virology,* 78**,** 8191-8200.

BAILLIE, J., SAHLENDER, D. & SINCLAIR, J. 2003. Human cytomegalovirus infection inhibits tumor necrosis factor alpha (TNF-α) signaling by targeting the 55-kilodalton TNF-α receptor. *Journal of virology,* 77**,** 7007-7016.

BALDICK, C. & SHENK, T. 1996. Proteins associated with purified human cytomegalovirus particles. *Journal of virology,* 70**,** 6097-6105.

BARRON, M. A., GAO, D., SPRINGER, K. L., PATTERSON, J. A., BRUNVAND, M. W., MCSWEENEY, P. A., CHANG, Z., BARÓN, A. E. & WEINBERG, A. 2009. Relationship of reconstituted adaptive and innate cytomegalovirus (CMV)-specific immune responses with CMV viremia in hematopoietic stem cell transplant recipients. *Clinical infectious diseases,* 49**,** 1777.

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.

BAUMGARTNER, M., SCHNEIDER, R., AUER, B., HERZOG, H., SCHWEIGER, M. & HIRSCH-KAUFFMANN, M. 1992. Fluorescence in situ mapping of the human nuclear NAD^+ ADP-ribosyltransferase gene (ADPRT) and two secondary sites to human chromosomal bands 1q42, 13q34, and 14q24. *Cytogenetic and Genome Research,* 61**,** 172-174.

BAXTER, M. & GIBSON, W. The putative human cytomegalovirus triplex proteins, minor capsid protein (mCP) and mCP-binding protein (mC-BP), form a heterotrimeric complex that localizes to the cell nucleus in the absence of other viral proteins. 1997.

BECHTEL, J. T. & SHENK, T. 2002. Human cytomegalovirus UL47 tegument protein functions after entry and before immediate-early gene expression. *Journal of virology,* 76**,** 1043-1050.

BECK, S. & BARRELL, B. G. 1988. Human cytomegalovirus encodes a glycoprotein homologous to MHC class-I antigens. *Nature,* 331**,** 269-272.

BEGO, M., MACIEJEWSKI, J., KHAIBOULLINA, S., PARI, G. & ST JEOR, S. 2005. Characterization of an antisense transcript spanning the UL81-82 locus of human cytomegalovirus. *Journal of virology,* 79**,** 11022.

BEGO, M. G., KEYES, L. R., MACIEJEWSKI, J. & ST JEOR, S. C. 2011. Human cytomegalovirus latency-associated protein LUNA is expressed during HCMV infections in vivo. *Archives of virology***,** 1-5.

BENEDICT, C. A., BUTROVICH, K. D., LURAIN, N. S., CORBEIL, J., ROONEY, I., SCHNEIDER, P., TSCHOPP, J. & WARE, C. F. 1999. Cutting edge: a novel viral TNF receptor superfamily member in virulent strains of human cytomegalovirus. *The Journal of Immunology,* 162**,** 6967-6970.

BERGAMINI, G., RESCHKE, M., BATTISTA, M. C., BOCCUNI, M. C., CAMPANINI, F., RIPALTI, A. & LANDINI, M. P. 1998. The major open reading frame of the β2. 7 transcript of human cytomegalovirus: in vitro expression of a protein posttranscriptionally regulated by the 5′ region. *Journal of virology,* 72**,** 8425-8429.

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.

BEWIG, B. & SCHMIDT, W. 2000. Accelerated titering of adenoviruses. *Biotechniques,* 28**,** 870.

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 11. *Journal of molecular biology,* 295**,** 155-161.

BIÈCHE, I., CHAMPÈME, M. H. & LIDEREAU, R. 1995. Loss and gain of distinct regions of chromosome 1q in primary breast cancer. *Clinical cancer research,* 1**,** 123-127.

BIEGALKE, B., LESTER, E., BRANDA, A. & RANA, R. 2004. Characterization of the human cytomegalovirus UL34 gene. *Journal of virology,* 78**,** 9579-9583.

BIRON, C. A., BYRON, K. S. & SULLIVAN, J. L. 1989. Severe herpesvirus infections in an adolescent without natural killer cells. *New England Journal of Medicine,* 320**,** 1731-1735.

BLANKENSHIP, C. A. & SHENK, T. 2002. Mutant human cytomegalovirus lacking the immediate-early TRS1 coding region exhibits a late defect. *Journal of virology,* 76**,** 12290-12299.

BOECKH, M. 2011. Complications, Diagnosis, Management, and Prevention of CMV Infections: Current and Future. *ASH Education Program Book,* 2011**,** 305-309.

BOECKH, M. & GEBALLE, A. P. 2011. Cytomegalovirus: pathogen, paradigm, and puzzle. *The Journal of clinical investigation,* 121**,** 1673.

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.

BOLOVAN-FRITTS, C. A., MOCARSKI, E. S. & WIEDEMAN, J. A. 1999. Peripheral blood CD14+ cells from healthy subjects carry a circular conformation of latent cytomegalovirus genome. *Blood,* 93**,** 394-398.

BORST, E. M., WAGNER, K., BINZ, A., SODEIK, B. & MESSERLE, M. 2008. The essential human cytomegalovirus gene UL52 is required for cleavage-packaging of the viral genome. *Journal of virology,* 82**,** 2065-2078.

BOSHART, M., WEBER, F., JAHN, G., DORSCH-HÄSLER, K., FLECKENSTEIN, B. & SCHAFFNER, W. 1985. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus.

BOTTINO, C., CASTRICONI, R., PENDE, D., RIVERA, P., NANNI, M., CARNEMOLLA, B., CANTONI, C., GRASSI, J., MARCENARO, S. & REYMOND, N. 2003. Identification of PVR (CD155) and Nectin-2 (CD112) as cell surface ligands for the human DNAM-1 (CD226) activating molecule. *The Journal of experimental medicine,* 198**,** 557-567.

BOWMAN, J. J., LACAYO, J. C., BURBELO, P., FISCHER, E. R. & COHEN, J. I. 2011. Rhesus and human cytomegalovirus glycoprotein L are required for infection and cell-to-cell spread of virus but cannot complement each other. *Journal of virology,* 85**,** 2089-2099.

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.

BRADSHAW, P. A., DURAN-GUARINO, M. R., PERKINS, S., ROWE, J. I., FERNANDEZ, J., FRY, K. E., REYES, G. R., YOUNG, L. & FOUNG, S. K. H. 1994. Localization of antigenic sites on human cytomegalovirus virion structural proteins encoded by UL48 and UL56. *Virology,* 205**,** 321-328.

BRAUD, V., JONES, E. Y. & MCMICHAEL, A. 1997. The human major histocompatibility complex class Ib molecule HLA-E binds signal sequence-derived peptides with primary anchor residues at positions 2 and 9. *European journal of immunology,* 27**,** 1164-9.

BRAUD, V. M., ALLAN, D. S., O'CALLAGHAN, C. A., SODERSTROM, K., D'ANDREA, A., OGG, G. S., LAZETIC, S., YOUNG, N. T., BELL, J. I., PHILLIPS, J. H., LANIER, L. L. & MCMICHAEL, A. J. 1998. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature,* 391**,** 795-9.

BRESNAHAN, W. A. & SHENK, T. 2000a. A subset of viral transcripts packaged within human cytomegalovirus particles. *Science,* 288**,** 2373-2376.

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

BRITT, B. 2007. Maturation and egress. *Human herpesviruses: biology, therapy, and immunoprophylaxis. Cambridge University Press, Cambridge, United Kingdom.*

BRITT, W. J. & AUGER, D. 1986. Synthesis and processing of the envelope gp55-116 complex of human cytomegalovirus. *Journal of virology,* 58**,** 185-191.

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

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.

BROWN, H. L. & ABERNATHY, M. P. Cytomegalovirus infection. 1998. Elsevier, 260-266.

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.

BROWNE, E. P., WING, B., COLEMAN, D. & SHENK, T. 2001. Altered cellular mRNA levels in human cytomegalovirus-infected fibroblasts: viral block to the accumulation of antiviral mRNAs. *Journal of virology,* 75**,** 12319-12330.

BRUNE, W. 2011. Inhibition of programmed cell death by cytomegaloviruses. *Virus Research,* 157**,** 144-150.

BRYANT, L., MIXON, P., DAVIDSON, M., BANNISTER, A., KOUZARIDES, T. & SINCLAIR, J. 2000. The human cytomegalovirus 86-kilodalton major immediate-early protein interacts physically and functionally with histone acetyltransferase P/CAF. *Journal of virology,* 74**,** 7230-7237.

BUERGER, I., REEFSCHLAEGER, J., BENDER, W., ECKENBERG, P., POPP, A., WEBER, O., GRAEPER, S., KLENK, H. D., RUEBSAMEN-WAIGMANN, H. & HALLENBERGER, S. 2001. A novel nonnucleoside inhibitor specifically targets cytomegalovirus DNA maturation via the UL89 and UL56 gene products. *Journal of virology,* 75**,** 9077-9086.

CALIENDO, A. M., GEORGE, K. S., KAO, S. Y., ALLEGA, J., TAN, B. H., LAFONTAINE, R., BUI, L. & RINALDO, C. R. 2000. Comparison of quantitative cytomegalovirus (CMV) PCR in plasma and CMV antigenemia assay: clinical utility of the prototype AMPLICOR CMV MONITOR test in transplant recipients. *Journal of clinical microbiology,* 38**,** 2122-2127.

CANNON, M. J., SCHMID, D. S. & HYDE, T. B. 2010. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Reviews in medical virology,* 20**,** 202-213.

CAO, J. & GEBALLE, A. P. 1995. Translational inhibition by a human cytomegalovirus upstream open reading frame despite inefficient utilization of its AUG codon. *Journal of virology,* 69**,** 1030-1036.

CAO, J. & GEBALLE, A. P. 1996. Coding sequence-dependent ribosomal arrest at termination of translation. *Molecular and Cellular Biology,* 16**,** 603-608.

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**,** 264-276.

CASWELL, R., BRYANT, L. & SINCLAIR, J. 1996. Human cytomegalovirus immediate-early 2 (IE2) protein can transactivate the human hsp70 promoter by alleviation of Dr1-mediated repression. *Journal of virology,* 70**,** 4028-4037.

CASWELL, R., HAGEMEIER, C., CHIOU, C. J., HAYWARD, G., KOUZARIDES, T. & SINCLAIR, J. 1993. The human cytomegalovirus 86K immediate early (IE) 2 protein requires the basic region of the TATA-box binding protein (TBP) for binding, and interacts with TBP and transcription factor TFIIB via regions of IE2 required for transcriptional regulation. *The Journal of general virology,* 74**,** 2691.

CERBONI, C., MOUSAVI-JAZI, M., LINDE, A., SÖDERSTRÖM, K., BRYTTING, M., WAHREN, B., KÄRRE, K. & CARBONE, E. 2000. Human cytomegalovirus strain-dependent changes in NK cell recognition of infected fibroblasts. *The Journal of Immunology,* 164**,** 4775.

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. *Journal of virology,* 70**,** 78.

CHALUPNY, N. J., REIN-WESTON, A., DOSCH, S. & COSMAN, D. 2006. Down-regulation of the NKG2D ligand MICA by the human cytomegalovirus glycoprotein UL142. *Biochemical and biophysical research communications,* 346**,** 175-181.

CHAMPSAUR, M. & LANIER, L. L. 2010. Effect of NKG2D ligand expression on host immune responses. *Immunological reviews,* 235**,** 267-285.

CHAN, G., NOGALSKI, M. T. & YUROCHKO, A. D. 2009. Activation of EGFR on monocytes is required for human cytomegalovirus entry and mediates cellular motility. *Proceedings of the National Academy of Sciences,* 106**,** 22369-22374.

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

CHAUDHURI, A. R., JEOR, S. S. & MACIEJEWSKI, J. P. 1999. Apoptosis induced by human cytomegalovirus infection can be enhanced by cytokines to limit the spread of virus. *Experimental hematology,* 27**,** 1194-1203.

CHEE, M., BANKIER, A., BECK, S., BOHNI, R., BROWN, C., CERNY, R., HORSNELL, T., HUTCHISON 3RD, C., KOUZARIDES, T. & MARTIGNETTI, J. 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Current topics in microbiology and immunology,* 154**,** 125.

CHEN, A., KAO, Y. & BROWN, C. M. 2005. Translation of the first upstream ORF in the hepatitis B virus pregenomic RNA modulates translation at the core and polymerase initiation codons. *Nucleic acids research,* 33**,** 1169-1181.

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.

CHERRINGTON, J. M. & MOCARSKI, E. S. 1989. Human cytomegalovirus ie1 transactivates the alpha promoter-enhancer via an 18-base-pair repeat element. *Journal of virology,* 63**,** 1435-1440.

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.

CHOU, S. 1986. Acquisition of donor strains of cytomegalovirus by renal-transplant recipients. *New England Journal of Medicine,* 314**,** 1418-1423.

CHOU, S. 1999. Antiviral drug resistance in human cytomegalovirus. *Transplant infectious disease,* 1**,** 105-114.

CHOU, S., MAROUSEK, G. I., SENTERS, A. E., DAVIS, M. G. & BIRON, K. K. 2004. Mutations in the human cytomegalovirus UL27 gene that confer resistance to maribavir. *Journal of virology,* 78**,** 7124-7130.

CHRETIEN, J. H., MCGINNISS, C. G. & MULLER, A. 1977. Venereal causes of cytomegalovirus mononucleosis. *JAMA: The Journal of the American Medical Association,* 238**,** 1644-1645.

CINATL JR, J., CINATL, J., VOGEL, J. U., RABENAU, H., KORNHUBER, B. & DOERR, H. 1996. Modulatory effects of human cytomegalovirus infection on malignant properties of cancer cells. *Intervirology,* 39**,** 259-269.

COBBS, C. S., HARKINS, L., SAMANTA, M., GILLESPIE, G. Y., BHARARA, S., KING, P. H., NABORS, L. B., COBBS, C. G. & BRITT, W. J. 2002. Human cytomegalovirus infection and expression in human malignant glioma. *Cancer research,* 62**,** 3347.

COLBERG-POLEY, A. M., PATEL, M. B., EREZO, D. P. P. & SLATER, J. E. 2000. Human cytomegalovirus UL37 immediate-early regulatory proteins traffic through the secretory apparatus and to mitochondria. *Journal of general virology,* 81**,** 1779-1789.

COMPTON, T., KURT-JONES, E. A., BOEHME, K. W., BELKO, J., LATZ, E., GOLENBOCK, D. T. & FINBERG, R. W. 2003. Human cytomegalovirus activates inflammatory cytokine responses via CD14 and Toll-like receptor 2. *Journal of virology,* 77**,** 4588-4596.

COMPTON, T., NEPOMUCENO, R. R. & NOWLIN, D. M. 1992. Human cytomegalovirus penetrates host cells by pH-independent fusion at the cell surface. *Virology,* 191**,** 387-395.

COMPTON, T., NOWLIN, D. M. & COOPER, N. R. 1993. Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate. *Virology,* 193**,** 834-841.

COSMAN, D., FANGER, N., BORGES, L., KUBIN, M., CHIN, W., PETERSON, L. & HSU, M. L. 1997. A novel immunoglobulin superfamily receptor for cellular and viral MHC class I molecules. *Immunity,* 7**,** 273-282.

CRANAGE, M., KOUZARIDES, T., BANKIER, A., SATCHWELL, S., WESTON, K., TOMLINSON, P., BARRELL, B., HART, H., BELL, S. & MINSON, A. 1986. Identification of the human cytomegalovirus glycoprotein B gene and induction of neutralizing antibodies via its expression in recombinant vaccinia virus. *The EMBO journal,* 5**,** 3057.

CRANAGE, M. P., SMITH, G., BELL, S., HART, H., BROWN, C., BANKIER, A., TOMLINSON, P., BARRELL, B. & MINSON, T. 1988. Identification and expression of a human cytomegalovirus glycoprotein with homology to the Epstein-Barr virus BXLF2 product, varicella-zoster virus gpIII, and herpes simplex virus type 1 glycoprotein H. *Journal of virology,* 62**,** 1416-1422.

CUNHA, B. A. 2010. Cytomegalovirus pneumonia: community-acquired pneumonia in immunocompetent hosts. *Infectious disease clinics of North America,* 24**,** 147-158.

DAL MONTE, P., PIGNATELLI, S., MACH, M. & LANDINI, M. 2001. The product of human cytomegalovirus UL73 is a new polymorphic structural glycoprotein (gpUL73). *Journal of human virology,* 4**,** 26.

DAL MONTE, P., PIGNATELLI, S., ZINI, N., MARALDI, N., PERRET, E., PREVOST, M. & LANDINI, M. 2002. Analysis of intracellular and intraviral localization of the human cytomegalovirus UL53 protein. *Journal of general virology,* 83**,** 1005-1012.

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., 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. *Journal of virology,* 81**,** 11861-11869.

DAVISON, A. & BHELLA, D. 2007. Comparative genome and virion structure. *Human herpesviruses: biology, therapy, and immunoprophylaxis. Cambridge University Press, Cambridge, United Kingdom.*

DAVISON, A. J. 2010. Herpesvirus systematics. *Veterinary microbiology,* 143**,** 52-69.

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.

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. *Journal of General Virology,* 84**,** 17-28.

DEGNIN, C. R., SCHLEISS, M., CAO, J. & GEBALLE, A. P. 1993. Translational inhibition mediated by a short upstream open reading frame in the human cytomegalovirus gpUL4 (gp48) transcript. *Journal of virology,* 67**,** 5514-5521.

DEMARCHI, J. M., SCHMIDT, C. A. & KAPLAN, A. S. 1980. Patterns of transcription of human cytomegalovirus in permissively infected cells. *Journal of virology,* 35**,** 277-286.

DEPTO, A. S. & STENBERG, R. M. 1992. Functional analysis of the true late human cytomegalovirus pp28 upstream promoter: cis-acting elements and viral trans-acting proteins necessary for promoter activation. *Journal of virology,* 66**,** 3241-3246.

DHURUVASAN, K., SIVASUBRAMANIAN, G. & PELLETT, P. E. 2011. Roles of host and viral microRNAs in human cytomegalovirus biology. *Virus Research,* 157**,** 180-192.

DIMITROPOULOU, P., CASWELL, R., MCSHARRY, B. P., GREAVES, R. F., SPANDIDOS, D. A., WILKINSON, G. W. G. & SOURVINOS, G. 2010. Differential relocation and stability of PML-body components during productive human cytomegalovirus infection: Detailed characterization by live-cell imaging. *European journal of cell biology,* 89**,** 757-768.

DITTMER, A. & BOGNER, E. 2005. Analysis of the quaternary structure of the putative HCMV portal protein pUL104. *Biochemistry,* 44**,** 759-765.

DIVANOVIC, S., TROMPETTE, A., ATABANI, S. F., MADAN, R., GOLENBOCK, D. T., VISINTIN, A., FINBERG, R. W., TARAKHOVSKY, A., VOGEL, S. N. & BELKAID, Y. 2005. Negative regulation of Toll-like receptor 4 signaling by the Toll-like receptor homolog RP105. *Nature immunology,* 6**,** 571-578.

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. 2004. Genetic content of wild-type human cytomegalovirus. *Journal of General Virology,* 85**,** 1301.

DÖLKEN, L., PFEFFER, S. & KOSZINOWSKI, U. H. 2009. Cytomegalovirus microRNAs. *Virus genes,* 38**,** 355-364.

DUCANCELLE, A., CHAMPIER, G., ALAIN, S., PETIT, F., LE PORS, M. J. S. & MAZERON, M. C. 2006. Short communication A novel mutation in the UL54 gene of human cytomegalovirus isolates that confers resistance to foscarnet. *Antiviral therapy,* 11**,** 537-540.

DUMORTIER, J., STREBLOW, D. N., MOSES, A. V., JACOBS, J. M., KREKLYWICH, C. N., CAMP, D., SMITH, R. D., ORLOFF, S. L. & NELSON, J. A. 2008. Human cytomegalovirus secretome contains factors that induce angiogenesis and wound healing. *Journal of virology,* 82**,** 6524-6535.

DUNN, C., CHALUPNY, N. J., SUTHERLAND, C. L., DOSCH, S., SIVAKUMAR, P., 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. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 14223.

DUNN, W., CHOU, C., LI, H., HAI, R., PATTERSON, D., STOLC, V., ZHU, H. & LIU, F. 2003c. Functional profiling of a human cytomegalovirus genome. *Proceedings of the National Academy of Sciences,* 100**,** 14223-14228.

DUNN, W., TRANG, P., ZHONG, Q., YANG, E., VAN BELLE, C. & LIU, F. 2005. Human cytomegalovirus expresses novel microRNAs during productive viral infection. *Cellular microbiology,* 7**,** 1684-1695.

E. MURPHY , T. S. 2008. *human cytomegalovirus* Current Topics in Microbiology and Immunology.

EDWARDS, R. L. 1976. Structural proteins of human cytomegalovirus. *The Yale journal of biology and medicine,* 49**,** 65.

ELSING, A. & BURGERT, H. G. 1998. The adenovirus E3/10.4 K–14.5 K proteins down-modulate the apoptosis receptor Fas/Apo-1 by inducing its internalization. *Proceedings of the National Academy of Sciences,* 95**,** 10072.

ENGEL, P., PÉREZ-CARMONA, N., ALBÀ, M. M., ROBERTSON, K., GHAZAL, P. & ANGULO, A. 2011. Human cytomegalovirus UL7, a homologue of the SLAM-family receptor CD229, impairs cytokine production. *Immunology and Cell Biology*.

ERTL, P. & POWELL, K. 1992. Physical and functional interaction of human cytomegalovirus DNA polymerase and its accessory protein (ICP36) expressed in insect cells. *Journal of virology,* 66**,** 4126-4133.

FARBER, S. & WOLBACH, S. B. 1932. Intranuclear and cytoplasmic inclusions ("protozoan-like bodies") in the salivary glands and other organs of infants. *The American journal of pathology,* 8**,** 123.

FEHR, A. R. & YU, D. 2010. Human cytomegalovirus gene UL21a encodes a short-lived cytoplasmic protein and facilitates virus replication in fibroblasts. *Journal of virology,* 84**,** 291-302.

FEIRE, A. L., KOSS, H. & COMPTON, T. 2004. Cellular integrins function as entry receptors for human cytomegalovirus via a highly conserved disintegrin-like domain. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 15470.

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.

FIALA, M., HONESS, R., HEINER, D., HEINE JR, J., MURNANE, J., WALLACE, R. & GUZE, L. 1976. Cytomegalovirus proteins. I. Polypeptides of virions and dense bodies. *Journal of virology,* 19**,** 243-254.

FISH, K. N., DEPTO, A. S., MOSES, A. V., BRITT, W. & NELSON, J. A. 1995. Growth kinetics of human cytomegalovirus are altered in monocyte-derived macrophages. *Journal of virology,* 69**,** 3737.

FLECKENSTEIN, B., MÜLLER, I. & COLLINS, J. 1982. Cloning of the complete human cytomegalovirus genome in cosmids. *Gene,* 18**,** 39-46.

FORTUNATO, E. A., DELL'AQUILA, M. L. & SPECTOR, D. H. 2000. Specific chromosome 1 breaks induced by human cytomegalovirus. *Proceedings of the National Academy of Sciences,* 97**,** 853.

FORTUNATO, E. A. & SPECTOR, D. H. 2003. Viral induction of site-specific chromosome damage. *Reviews in medical virology,* 13**,** 21-37.

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.

FRAILE-RAMOS, A., KLEDAL, T. N., PELCHEN-MATTHEWS, A., BOWERS, K., SCHWARTZ, T. W. & MARSH, M. 2001. The human cytomegalovirus US28 protein is located in endocytic vesicles and undergoes constitutive endocytosis and recycling. *Molecular biology of the cell,* 12**,** 1737-1749.

FRAILE-RAMOS, A., PELCHEN-MATTHEWS, A., KLEDAL, T. N., BROWNE, H., SCHWARTZ, T. W. & MARSH, M. 2002. Localization of HCMV UL33 and US27 in endocytic compartments and viral membranes. *Traffic,* 3**,** 218-232.

FREIBERG, R. A., SPENCER, D. M., CHOATE, K. A., DUH, H. J., SCHREIBER, S. L., CRABTREE, G. R. & KHAVARI, P. A. 1997. Fas signal transduction triggers either proliferation or apoptosis in human fibroblasts. *Journal of investigative dermatology,* 108**,** 215-219.

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.

FURNARI, B. A., POMA, E., KOWALIK, T., HUONG, S. & HUANG, E. 1993. Human cytomegalovirus immediate-early gene 2 protein interacts with itself and with several novel cellular proteins. *Journal of virology,* 67**,** 4981-4991.

FURUKAWA, T., HORNBERGER, E., SAKUMA, S. & PLOTKIN, S. A. 1975. Demonstration of immunoglobulin G receptors induced by human cytomegalovirus. *Journal of clinical microbiology,* 2**,** 332-336.

GANDHI, M. K. & KHANNA, R. 2004. Human cytomegalovirus: clinical aspects, immune regulation, and emerging treatments. *The Lancet infectious diseases,* 4**,** 725-738.

GAO, M., BRUFATTO, N., CHEN, T., MURLEY, L. L., THALAKADA, R., DOMAGALA, M., BEATTIE, B., MAMELAK, D., ATHANASOPOULOS, V. & JOHNSON, D. 2005. Expression profiling of herpesvirus and vaccinia virus proteins using a high-throughput baculovirus screening system. *Journal of Proteome Research,* 4**,** 2225-2235.

GARCIA-FUSTER, M. J., FERRER-ALCÓN, M., MIRALLES, A. & GARCíA-SEVILLA, J. A. 2004. Deglycosylation of Fas receptor and chronic morphine treatment up-regulate high molecular mass Fas aggregates in the rat brain. *European journal of pharmacology,* 496**,** 63-69.

GATHERER, D., SEIRAFIAN, S., CUNNINGHAM, C., HOLTON, M., DARGAN, D. J., BALUCHOVA, K., HECTOR, R. D., GALBRAITH, J., HERZYK, P. & WILKINSON, G. W. G. 2011. High-resolution human cytomegalovirus transcriptome. *Proceedings of the National Academy of Sciences,* 108**,** 19755-19760.

GAWN, J. M. & GREAVES, R. F. 2002. Absence of IE1 p72 protein function during low-multiplicity infection by human cytomegalovirus results in a broad block to viral delayed-early gene expression. *Journal of virology,* 76**,** 4441-4455.

GEBALLE, A. P., LEACH, F. S. & MOCARSKI, E. S. 1986. Regulation of cytomegalovirus late gene expression: gamma genes are controlled by posttranscriptional events. *Journal of virology,* 57**,** 864-874.

GEDER, K., LAUSCH, R., O'NEILL, F. & RAPP, F. 1976. Oncogenic transformation of human embryo lung cells by human cytomegalovirus. *Science,* 192**,** 1134-1137.

GEORGEL, P., JIANG, Z., KUNZ, S., JANSSEN, E., MOLS, J., HOEBE, K., BAHRAM, S., OLDSTONE, M. & BEUTLER, B. 2007. Vesicular stomatitis virus glycoprotein G activates a specific antiviral Toll-like receptor 4-dependent pathway. *Virology,* 362**,** 304-313.

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

GERNA, G., SARASINI, A., PATRONE, M., PERCIVALLE, E., FIORINA, L., CAMPANINI, G., GALLINA, A., BALDANTI, F. & REVELLO, M. G. 2008b. Human cytomegalovirus serum neutralizing antibodies block virus infection of endothelial/epithelial cells, but not fibroblasts, early during primary infection. *Journal of general virology,* 89**,** 853-865.

GHAZAL, P., LUBON, H., REYNOLDS-KOHLER, C., HENNIGHAUSEN, L. & NELSON, J. A. 1990. Interactions between cellular regulatory proteins and a unique sequence region in the human cytomegalovirus major immediate-early promoter. *Virology,* 174**,** 18-25.

GIBSON, W. 1981. Structural and nonstructural proteins of strain Colburn cytomegalovi. *Virology,* 111**,** 516-537.

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

GIBSON, W., BAXTER, M. K. & CLOPPER, K. S. 1996a. Cytomegalovirus" missing" capsid protein identified as heat-aggregable product of human cytomegalovirus UL46. *Journal of virology,* 70**,** 7454-7461.

GIBSON, W., CLOPPER, K. S., BRITT, W. J. & BAXTER, M. K. 1996. Human cytomegalovirus (HCMV) smallest capsid protein identified as product of short open reading frame located between HCMV UL48 and UL49. *Journal of virology,* 70**,** 5680-5683.

GIBSON, W. & ROIZMAN, B. 1971. Compartmentalization of spermine and spermidine in the herpes simplex virion. *Proceedings of the National Academy of Sciences,* 68**,** 2818.

GIESEN, K., RADSAK, K. & BOGNER, E. 2000. Targeting of the gene product encoded by ORF UL56 of human cytomegalovirus into viral replication centers. *FEBS letters,* 471**,** 215-218.

GLAß, M., BUSCHE, A., WAGNER, K., MESSERLE, M. & BORST, E. M. 2009. Conditional and reversible disruption of essential herpesvirus proteins. *Nature methods,* 6**,** 577-579.

GOLDMACHER, V. 2005. Cell death suppression by cytomegaloviruses. *Apoptosis,* 10**,** 251-265.

GOLDMACHER, V. S., BARTLE, L. M., SKALETSKAYA, A., DIONNE, C. A., KEDERSHA, N. L., VATER, C. A., HAN, J., LUTZ, R. J., WATANABE, S. & MCFARLAND, E. D. C. 1999. A cytomegalovirus-encoded mitochondria-localized inhibitor of apoptosis structurally unrelated to Bcl-2. *Proceedings of the National Academy of Sciences,* 96**,** 12536.

GOLDSTEIN, L., MCDOUGALL, J., HACKMAN, R., MEYERS, J., THOMAS, E. & NOWINSKI, R. 1982. Monoclonal antibodies to cytomegalovirus: rapid identification of clinical isolates and preliminary use in diagnosis of cytomegalovirus pneumonia. *Infection and immunity,* 38**,** 273-281.

GONCZOL, E., ANDREWS, P. W. & PLOTKIN, S. A. 1984. Cytomegalovirus replicates in differentiated but not in undifferentiated human embryonal carcinoma cells. *Science,* 224**,** 159-161.

GOODPASTURE, E. W. & TALBOT, F. B. 1921. Concerning the nature of" protozoan-like" cells in certain lesions of infancy. *Archives of Pediatrics and Adolescent Medicine,* 21**,** 415.

GOODRUM, F., REEVES, M., SINCLAIR, J., HIGH, K. & SHENK, T. 2007. Human cytomegalovirus sequences expressed in latently infected individuals promote a latent infection in vitro. *Blood,* 110**,** 937-945.

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. *Proceedings of the National Academy of Sciences,* 99**,** 16255.

GRAHAM, F., SMILEY, J., RUSSELL, W. & NAIRN, R. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *Journal of general virology,* 36**,** 59-72.

GRAINGER, L., CICCHINI, L., RAK, M., PETRUCELLI, A., FITZGERALD, K. D., SEMLER, B. L. & GOODRUM, F. 2010. Stress-inducible alternative translation initiation of human cytomegalovirus latency protein pUL138. *Journal of virology,* 84**,** 9472-9486.

GRAZIA REVELLO, M., BALDANTI, F., PERCIVALLE, E., SARASINI, A., DE-GIULI, L., GENINI, E., LILLERI, D., LABÒ, N. & GERNA, G. 2001. In vitro selection of human cytomegalovirus variants unable to transfer virus and virus products from infected cells to polymorphonuclear leukocytes and to grow in endothelial cells. *Journal of General Virology,* 82**,** 1429.

GREAVES, R. F. & MOCARSKI, E. S. 1998. Defective growth correlates with reduced accumulation of a viral DNA replication protein after low-multiplicity infection by a human cytomegalovirus ie1 mutant. *Journal of virology,* 72**,** 366-379.

GREBER, U. F. & WAY, M. 2006. A superhighway to virus infection. *Cell,* 124**,** 741-754.

GREENAWAY, P. J. & WILKINSON, G. W. G. 1987. Nucleotide sequence of the most abundantly transcribed early gene of human cytomegalovirus strain AD169. *Virus Research,* 7**,** 17-31.

GREFTE, A., HARMSEN, M. C., VAN DER GIESSEN, M., KNOLLEMA, S. & VAN SON, W. J. 1994. Presence of human cytomegalovirus (HCMV) immediate early mRNA but not ppUL83 (lower matrix protein pp65) mRNA in polymorphonuclear and mononuclear leukocytes during active HCMV infection. *The Journal of general virology,* 75**,** 1989.

GREIJER, A. E., DEKKERS, C. A. J. & MIDDELDORP, J. M. 2000. Human cytomegalovirus virions differentially incorporate viral and host cell RNA during the assembly process. *Journal of virology,* 74**,** 9078-9082.

GREY, F., ANTONIEWICZ, A., ALLEN, E., SAUGSTAD, J., MCSHEA, A., CARRINGTON, J. C. & NELSON, J. 2005. Identification and characterization of human cytomegalovirus-encoded microRNAs. *Journal of virology,* 79**,** 12095-12099.

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. *Journal of General Virology,* 86**,** 2999-3008.

GRUNDY, J. E., MCKEATING, J. A. & GRIFFITHS, P. D. 1987. Cytomegalovirus strain AD169 binds beta 2 microglobulin in vitro after release from cells. *The Journal of general virology,* 68**,** 777.

GUERIN, J. L., GELFI, J., BOULLIER, S., DELVERDIER, M., BELLANGER, F. A., BERTAGNOLI, S., DREXLER, I., SUTTER, G. & MESSUD-PETIT, F. 2002. Myxoma virus leukemia-associated protein is responsible for major histocompatibility complex class I and Fas-CD95 down-regulation and defines scrapins, a new group of surface cellular receptor abductor proteins. *Journal of virology,* 76**,** 2912-2923.

GUO, Y. & HUANG, E. 1993. Characterization of a structurally tricistronic gene of human cytomegalovirus composed of U (s) 18, U (s) 19, and U (s) 20. *Journal of virology,* 67**,** 2043-2054.

HAGEMEIER, C., CASWELL, R., HAYHURST, G., SINCLAIR, J. & KOUZARIDES, T. 1994. Functional interaction between the HCMV IE2 transactivator and the retinoblastoma protein. *The EMBO journal,* 13**,** 2897.

HAHN, G., KHAN, H., BALDANTI, F., KOSZINOWSKI, U. H., REVELLO, M. G. & GERNA, G. 2002. The human cytomegalovirus ribonucleotide reductase homolog UL45 is dispensable for growth in endothelial cells, as determined by a BAC-cloned clinical isolate of human cytomegalovirus with preserved wild-type characteristics. *Journal of virology,* 76**,** 9551.

HAHN, G., REVELLO, M. G., PATRONE, M., PERCIVALLE, E., CAMPANINI, G., SARASINI, A., WAGNER, M., GALLINA, A., MILANESI, G. & KOSZINOWSKI, U. 2004. Human cytomegalovirus UL131-128 genes are indispensable for virus growth in endothelial cells and virus transfer to leukocytes. *Journal of virology,* 78**,** 10023-10033.

HANDSFIELD, H. H., CHANDLER, S. H., CAINE, V. A., MEYERS, J. D., COREY, L., MEDEIROS, E. & MCDOUGALL, J. K. 1985. Cytomegalovirus infection in sex partners: evidence for sexual transmission. *Journal of Infectious Diseases,* 151**,** 344.

HAREL, N. Y. & ALWINE, J. C. 1998. Phosphorylation of the human cytomegalovirus 86-kilodalton immediate-early protein IE2. *Journal of virology,* 72**,** 5481-92.

HARKINS, L., VOLK, A. L., SAMANTA, M., MIKOLAENKO, I., BRITT, W. J., BLAND, K. I. & COBBS, C. S. 2002. Specific localisation of human cytomegalovirus nucleic acids and proteins in human colorectal cancer. *The Lancet,* 360**,** 1557-1563.

HAYHURST, G., BRYANT, L., CASWELL, R., WALKER, S. & SINCLAIR, J. 1995. CCAAT box-dependent activation of the TATA-less human DNA polymerase alpha promoter by the human cytomegalovirus 72-kilodalton major immediate-early protein. *Journal of virology,* 69**,** 182-188.

HE, T. C., ZHOU, S., DA COSTA, L. T., YU, J., KINZLER, K. W. & VOGELSTEIN, B. 1998. A simplified system for generating recombinant adenoviruses. *Proceedings of the National Academy of Sciences,* 95**,** 2509-2514.

HE, Y., XU, L. & HUANG, E. 1992. Characterization of human cytomegalovirus UL84 early gene and identification of its putative protein product. *Journal of virology,* 66**,** 1098-1108.

HEGDE, N. R., TOMAZIN, R. A., WISNER, T. W., DUNN, C., BONAME, J. M., LEWINSOHN, D. M. & JOHNSON, D. C. 2002. Inhibition of HLA-DR assembly, transport, and loading by human cytomegalovirus glycoprotein US3: a novel mechanism for evading major histocompatibility complex class II antigen presentation. *Journal of virology,* 76**,** 10929-10941.

HENGEL, H., FLOHR, T., HÄMMERLING, G. J., KOSZINOWSKI, U. H. & MOMBURG, F. 1996. Human cytomegalovirus inhibits peptide translocation into the endoplasmic reticulum for MHC class I assembly. *Journal of General Virology,* 77**,** 2287-2296.

HENGEL, H., KOOPMANN, J. O., FLOHR, T., MURANYI, W., GOULMY, E., HÄMMERLING, G. J., KOSZINOWSKI, U. H. & MOMBURG, F. 1997. A viral ER-resident glycoprotein inactivates the MHC-encoded peptide transporter. *Immunity,* 6**,** 623-632.

HENSEL, G., MEYER, H., GÄRTNER, S., BRAND, G. & KERN, H. F. 1995. Nuclear localization of the human cytomegalovirus tegument protein pp150 (ppUL32). *Journal of General Virology,* 76**,** 1591-1601.

HENSEL, G. M., MEYER, H. H., BUCHMANN, I., POMMEREHNE, D., SCHMOLKE, S., PLACHTER, B., RADSAK, K. & KERN, H. F. 1996. Intracellular localization and expression of the human cytomegalovirus matrix phosphoprotein pp71 (ppUL82): evidence for its translocation into the nucleus. *Journal of general virology,* 77**,** 3087-3097.

HERMISTON, T. W., MALONE, C. L. & STINSKI, M. F. 1990. Human cytomegalovirus immediate-early two protein region involved in negative regulation of the major immediate-early promoter. *Journal of virology,* 64**,** 3532-3536.

HITOMI, S., KOZUKA-HATA, H., CHEN, Z., SUGANO, S., YAMAGUCHI, N. & WATANABE, S. 1997. Human cytomegalovirus open reading frame UL11 encodes a highly polymorphic protein expressed on the infected cell surface. *Archives of virology,* 142**,** 1407-1427.

HO, M. 1990. Epidemiology of cytomegalovirus infections. *Review of Infectious Diseases,* 12**,** S701-S710.

HO, M., SUWANSIRIKUL, S., DOWLING, J. N., YOUNGBLOOD, L. A. & ARMSTRONG, J. A. 1975. The transplanted kidney as a source of cytomegalovirus infection. *New England Journal of Medicine,* 293**,** 1109-1112.

HOBOM, U., BRUNE, W., MESSERLE, M., HAHN, G. & KOSZINOWSKI, U. H. 2000. Fast screening procedures for random transposon libraries of cloned herpesvirus genomes: mutational analysis of human cytomegalovirus envelope glycoprotein genes. *Journal of virology,* 74**,** 7720-7729.

HSU, C. H., CHANG, M. D. T., TAI, K. Y., YANG, Y. T., WANG, P. S., CHEN, C. J., WANG, Y. H., LEE, S. C., WU, C. W. & JUAN, L. J. 2004. HCMV IE2-mediated inhibition of HAT activity downregulates p53 function. *The EMBO journal,* 23**,** 2269-2280.

HUBER, M. T. & COMPTON, T. 1998. The human cytomegalovirus UL74 gene encodes the third component of the glycoprotein H-glycoprotein L-containing envelope complex. *Journal of virology,* 72**,** 8191-8197.

HUBER, M. T. & COMPTON, T. 1999. Intracellular formation and processing of the heterotrimeric gH-gL-gO (gCIII) glycoprotein envelope complex of human cytomegalovirus. *Journal of virology,* 73**,** 3886-3892.

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.

HUTCHINSON, N. I. & TOCCI, M. J. 1986. Characterization of a major early gene from the human cytomegalovirus long inverted repeat; predicted amino acid sequence of a 30-kDa protein encoded by the 1.2-kb mRNA. *Virology,* 155**,** 172-182.

INABA, T., SHIMANO, H., GOTODA, T., HARADA, K., SHIMADA, M., OHSUGA, J., WATANABE, Y., KAWAMURA, M., YAZAKI, Y. & YAMADA, N. 1993. Expression of platelet-derived growth factor beta receptor on human monocyte-derived macrophages and effects of platelet-derived growth factor BB dimer on the cellular function. *Journal of Biological Chemistry,* 268**,** 24353-24360.

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**,** 118-133.

ISAACSON, M., JUCKEM, L. & COMPTON, T. 2008. Virus entry and innate immune activation. *Human Cytomegalovirus***,** 85-100.

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. *Journal of virology,* 83**,** 3891-3903.

ISAACSON, M. K., FEIRE, A. L. & COMPTON, T. 2007. Epidermal growth factor receptor is not required for human cytomegalovirus entry or signaling. *Journal of virology,* 81**,** 6241-6247.

ISOMURA, H., STINSKI, M. F., MURATA, T., YAMASHITA, Y., KANDA, T., TOYOKUNI, S. & TSURUMI, T. 2011. The human cytomegalovirus gene products essential for late viral gene expression assemble into prereplication complexes before viral DNA replication. *Journal of virology,* 85**,** 6629-6644.

IWASAKI, Y., FURUKAWA, T., PLOTKIN, S. & KOPROWSKI, H. 1973. Ultrastructural study on the sequence of human cytomegalovirus infection in human diploid cells. *Archives of virology,* 40**,** 311-324.

JACKSON, S. E., MASON, G. M. & WILLS, M. R. 2011. Human cytomegalovirus immunity and immune evasion. *Virus Research,* 157**,** 151-160.

JACOBSON, M. A. & MILLS, J. 1988. Serious Cytomegalovirus Disease in the Acquired Immunodeficiency Syndrome (AIDS) Clinical Findings, Diagnosis, and Treatment. *Annals of internal medicine,* 108**,** 585-594.

JELASKA, A. & KORN, J. H. 1998. Anti-fas induces apoptosis and proliferation in human dermal fibroblasts: Differences between foreskin and adult fibroblasts. *Journal of cellular physiology,* 175**,** 19-29.

JELCIC, I., REICHEL, J., SCHLUDE, C., TREUTLER, E., SINZGER, C. & STEINLE, A. 2011. The Polymorphic HCMV Glycoprotein UL20 Is Targeted for Lysosomal Degradation by Multiple Cytoplasmic Dileucine Motifs. *Traffic,* 12**,** 1444-1456.

JENKINS, C., ABENDROTH, A. & SLOBEDMAN, B. 2004. A novel viral transcript with homology to human interleukin-10 is expressed during latent human cytomegalovirus infection. *Journal of virology,* 78**,** 1440-1447.

JESIONEK, A. & KIOLEMENOGLOU, B. 1904. Uber einen Befund von Protozoenartigen gebilden in den Organen eines heriditarluetischen Fotus. *Munch Med Wochenschr,* 51**,** 1905-1907.

JONES, T. R., MUZITHRAS, V. P. & GLUZMAN, Y. 1991. Replacement mutagenesis of the human cytomegalovirus genome: US10 and US11 gene products are nonessential. *Journal of virology,* 65**,** 5860-5872.

JONES, T. R. & SUN, L. 1997. Human cytomegalovirus US2 destabilizes major histocompatibility complex class I heavy chains. *Journal of virology,* 71**,** 2970.

JONES, T. R., WIERTZ, E., SUN, L., FISH, K. N., NELSON, J. A. & PLOEGH, H. L. 1996. Human cytomegalovirus US3 impairs transport and maturation of major histocompatibility complex class I heavy chains. *Proceedings of the National Academy of Sciences,* 93**,** 11327.

JORDAN, M. C., ROUSSEAU, W. E., NOBLE, G. R., STEWART, J. A. & CHIN, T. D. Y. 1973. Association of cervical cytomegaloviruses with venereal disease. *New England Journal of Medicine,* 288**,** 932-934.

JUPP, R., HOFFMANN, S., STENBERG, R., NELSON, J. & GHAZAL, P. 1993. Human cytomegalovirus IE86 protein interacts with promoter-bound TATA-binding protein via a specific region distinct from the autorepression domain. *Journal of virology,* 67**,** 7539-7546.

KÄGI, D., VIGNAUX, F., LEDERMANN, B., BÜRKI, K., DEPRAETERE, V., NAGATA, S., HENGARTNER, H. & GOLSTEIN, P. 1994. Fas and perforin pathways as major mechanisms of T cell-mediated cytotoxicity. *Science (New York, NY),* 265**,** 528.

KALEJTA, R. F. 2008. Tegument proteins of human cytomegalovirus. *Microbiology and molecular biology reviews,* 72**,** 249-265.

KARRE, K., LJUNGGREN, H. G., PIONTEK, G. & KIESSLING, R. 1986. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature,* 319**,** 675-8.

KAYE, J., BROWNE, H., STOFFEL, M. & MINSON, T. 1992. The UL16 gene of human cytomegalovirus encodes a glycoprotein that is dispensable for growth in vitro. *Journal of virology,* 66**,** 6609-6615.

KEAY, S. & BALDWIN, B. 1995. Update on the 92.5 kDa putative HCMV fusion receptor. *Scandinavian journal of infectious diseases. Supplementum,* 99**,** 32.

KELLER, R., PEITCHEL, R., GOLDMAN, J. N. & GOLDMAN, M. 1976. An IgG-Fc receptor induced in cytomegalovirus-infected human fibroblasts. *The Journal of Immunology,* 116**,** 772-777.

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.

KEMBLE, G., MCCORMICK, A., PEREIRA, L. & MOCARSKI, E. 1987. A cytomegalovirus protein with properties of herpes simplex virus ICP8: partial purification of the polypeptide and map position of the gene. *Journal of virology,* 61**,** 3143-3151.

KEMBLE, G. W. & MOCARSKI, E. S. 1989. A host cell protein binds to a highly conserved sequence element (pac-2) within the cytomegalovirus a sequence. *Journal of virology,* 63**,** 4715.

KEPPLER, O. T., PETER, M. E., HINDERLICH, S., MOLDENHAUER, G., STEHLING, P., SCHMITZ, I., SCHWARTZ-ALBIEZ, R., REUTTER, W. & PAWLITA, M. 1999. Differential sialylation of cell surface glycoconjugates in a human B lymphoma cell line regulates susceptibility for CD95 (APO-1/Fas)-mediated apoptosis and for infection by a lymphotropic virus. *Glycobiology,* 9**,** 557-569.

KERRY, J. A., SEHGAL, A., BARLOW, S. W., CAVANAUGH, V. J., FISH, K., NELSON, J. A. & STENBERG, R. M. 1995. Isolation and characterization of a low-abundance splice variant from the human cytomegalovirus major immediate-early gene region. *Journal of virology,* 69**,** 3868-3872.

KHAN, N., COBBOLD, M., KEENAN, R. & MOSS, P. A. 2002. Comparative analysis of CD8+ T cell responses against human cytomegalovirus proteins pp65 and immediate early 1 shows

similarities in precursor frequency, oligoclonality, and phenotype. *The Journal of infectious diseases,* 185**,** 1025-34.
KILPATRICK, B. A., HUANG, E. & PAGANO, J. S. 1976. Analysis of cytomegalovirus genomes with restriction endonucleases Hin D III and EcoR-1. *Journal of virology,* 18**,** 1095-1105.
KIM, K., SAPIENZA, V., CARP, R. & MOON, H. 1976. Analysis of structural polypeptides of purified human cytomegalovirus. *Journal of virology,* 20**,** 604-611.
KOMAZIN, G., PTAK, R. G., EMMER, B. T., TOWNSEND, L. B. & DRACH, J. C. 2003. Resistance of human cytomegalovirus to the benzimidazole L-ribonucleoside maribavir maps to UL27. *Journal of virology,* 77**,** 11499-11506.
KOSLOWSKI, K. M., SHAVER, P. R., CASEY, J. T., WILSON, T., YAMANAKA, G., SHEAFFER, A. K., TENNEY, D. J. & PEDERSON, N. E. 1999. Physical and functional interactions between the herpes simplex virus UL15 and UL28 DNA cleavage and packaging proteins. *Journal of virology,* 73**,** 1704-1707.
KOTENKO, S. V., SACCANI, S., IZOTOVA, L. S., MIROCHNITCHENKO, O. V. & PESTKA, S. 2000a. Human cytomegalovirus harbors its own unique IL-10 homolog (cmvIL-10). *Proceedings of the National Academy of Sciences,* 97**,** 1695-1700.
KOTENKO, S. V., SACCANI, S., IZOTOVA, L. S., MIROCHNITCHENKO, O. V. & PESTKA, S. 2000b. Human cytomegalovirus harbors its own unique IL-10 homolog (cmvIL-10). *Proceedings of the National Academy of Sciences,* 97**,** 1695.
KOUZARIDES, T., BANKIER, A. T., SATCHWELL, S. C., PREDDY, E. & BARRELL, B. G. 1988. An immediate early gene of human cytomegalovirus encodes a potential membrane glycoprotein. *Virology,* 165**,** 151-164.
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.
KULESZA, C. A. & SHENK, T. 2004. Human cytomegalovirus 5-kilobase immediate-early RNA is a stable intron. *Journal of virology,* 78**,** 13182-13189.
KURT-JONES, E. A., POPOVA, L., KWINN, L., HAYNES, L. M., JONES, L. P., TRIPP, R. A., WALSH, E. E., FREEMAN, M. W., GOLENBOCK, D. T. & ANDERSON, L. J. 2000. Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. *Nature immunology,* 1**,** 398-401.
LAHIJANI, R. S., OTTESON, E. W. & ST JEOR, S. C. 1992. A possible role for nonsense suppression in the synthesis of a human cytomegalovirus 58-kDa virion protein. *Virology,* 186**,** 309-312.
LAI, L. & BRITT, W. J. 2003. The interaction between the major capsid protein and the smallest capsid protein of human cytomegalovirus is dependent on two linear sequences in the smallest capsid protein. *Journal of virology,* 77**,** 2730-2735.
LANDINI, M., SEVERI, B., FURLINI, G. & BADIALI DE GIORGI, L. 1987. Human cytomegalovims structural components: intracellular and intraviral localization of p28 and p65–69 by immunoelectron microscopy. *Virus Research,* 8**,** 15-23.
LANG, D., GEBERT, S., ARLT, H. & STAMMINGER, T. 1995. Functional interaction between the human cytomegalovirus 86-kilodalton IE2 protein and the cellular transcription factor CREB. *Journal of virology,* 69**,** 6030-6037.
LAPIERRE, L. A. & BIEGALKE, B. J. 2001. Identification of a novel transcriptional repressor encoded by human cytomegalovirus. *Journal of virology,* 75**,** 6062-6069.
LAWLOR, G. & MOSS, A. C. 2010. Cytomegalovirus in inflammatory bowel disease: pathogen or innocent bystander? *Inflammatory bowel diseases,* 16**,** 1620-1627.
LE, V. T. K., TRILLING, M. & HENGEL, H. 2011. The Cytomegaloviral Protein pUL138 Acts as Potentiator of Tumor Necrosis Factor (TNF) Receptor 1 Surface Density To Enhance ULb'-Encoded Modulation of TNF-α Signaling. *Journal of Virology,* 85**,** 13260-13270.
LEATHAM, M. P., WITTE, P. & STINSKI, M. F. 1991. Alternate promoter selection within a human cytomegalovirus immediate-early and early transcription unit (UL119-115)

defines true late transcripts containing open reading frames for putative viral glycoproteins. *Journal of virology,* 65**,** 6144-6153.

LEE, E., YU, D., MARTINEZ DE VELASCO, J., TESSAROLLO, L., SWING, D. A., JENKINS, N. A. & COPELAND, N. G. 2001. A Highly Efficient Escherichia coli-Based Chromosome Engineering System Adapted for Recombinogenic Targeting and Subcloning of BAC DNA. *Genomics,* 73**,** 56-65.

LEGUIZAMON, G. & REECE, E. 1997. Is serologic screening of all pregnant women for cytomegalovirus warranted? *CONTEMPORARY OB GYN,* 42**,** 49-62.

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. *Proceedings of the National Academy of Sciences of the United States of America,* 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,* 95**,** 13159.

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, Y., YANG, X., NGUYEN, A. H. T. & BROCKHAUSEN, I. 2007. Requirement of< i> N</i>-glycosylation for the secretion of recombinant extracellular domain of human Fas in HeLa cells. *The international journal of biochemistry & cell biology,* 39**,** 1625-1636.

LI, Y. S., RAMSAY, D. A., FAN, Y. S., ARMSTRONG, R. F. & DEL MAESTRO, R. F. 1995. Cytogenetic evidence that a tumor suppressor gene in the long arm of chromosome 1 contributes to glioma growth. *Cancer genetics and cytogenetics,* 84**,** 46-50.

LITTLER, E., STUART, A. D. & CHEE, M. S. 1992. Human cytomegalovirus UL97 open reading frame encodes a protein that phosphorylates the antiviral nucleoside analogue ganciclovir.

LIU, B. & STINSKI, M. F. 1992. Human cytomegalovirus contains a tegument protein that enhances transcription from promoters with upstream ATF and AP-1 cis-acting elements. *Journal of virology,* 66**,** 4434-4444.

LIU, F. & ZHOU, Z. H. 2007. Comparative virion structures of human herpesviruses. *Human herpesviruses: biology, therapy, and immunoprophylaxis. Cambridge University Press, Cambridge, United Kingdom.*

LIU, Y. & BIEGALKE, B. J. 2002. The human cytomegalovirus UL35 gene encodes two proteins with different functions. *Journal of virology,* 76**,** 2460-2468.

LIU, Y., CUI, Z., ZHANG, Z., WEI, H., ZHOU, Y., WANG, M. & ZHANG, X. E. 2009. The tegument protein UL94 of human cytomegalovirus as a binding partner for tegument protein pp28 identified by intracellular imaging. *Virology,* 388**,** 68-77.

LUKAC, D. M. & ALWINE, J. C. 1999. Effects of human cytomegalovirus major immediate-early proteins in controlling the cell cycle and inhibiting apoptosis: studies with ts13 cells. *Journal of virology,* 73**,** 2825-2831.

LUUKKONEN, B., TAN, W. & SCHWARTZ, S. 1995. Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *Journal of virology,* 69**,** 4086-4094.

MACAGNO, A., BERNASCONI, N. L., VANZETTA, F., DANDER, E., SARASINI, A., REVELLO, M. G., GERNA, G., SALLUSTO, F. & LANZAVECCHIA, A. 2010. Isolation of human monoclonal antibodies that potently neutralize human cytomegalovirus infection by targeting different epitopes on the gH/gL/UL128-131A complex. *Journal of virology,* 84**,** 1005-1013.

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.

MAIDJI, E., TUGIZOV, S., ABENES, G., JONES, T. & PEREIRA, L. 1998. A novel human cytomegalovirus glycoprotein, gpUS9, which promotes cell-to-cell spread in polarized epithelial cells, colocalizes with the cytoskeletal proteins E-cadherin and F-actin. *Journal of virology,* 72**,** 5717-5727.

MALONE, C. L., VESOLE, D. H. & STINSKI, M. F. 1990. Transactivation of a human cytomegalovirus early promoter by gene products from the immediate-early gene IE2 and augmentation by IE1: mutational analysis of the viral proteins. *Journal of virology,* 64**,** 1498-1506.

MARGOLIS, M. J., PAJOVIC, S., WONG, E. L., WADE, M., JUPP, R., NELSON, J. & AZIZKHAN, J. C. 1995. Interaction of the 72-kilodalton human cytomegalovirus IE1 gene product with E2F1 coincides with E2F-dependent activation of dihydrofolate reductase transcription. *Journal of virology,* 69**,** 7759-7767.

MARGULIES, B. J., BROWNE, H. & GIBSON, W. 1996. Identification of the human cytomegalovirus G protein-coupled receptor homologue encoded by UL33 in infected cells and enveloped virus particles. *Virology,* 225**,** 111-125.

MCGEOCH, D. J., DOLAN, A., DONALD, S. & RIXON, F. J. 1985. Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. *Journal of molecular biology,* 181**,** 1.

MCLAREN, J. E., CALDER, C. J., MCSHARRY, B. P., SEXTON, K., SALTER, R. C., SINGH, N. N., WILKINSON, G. W. G., WANG, E. C. Y. & RAMJI, D. P. 2010. The TNF-Like Protein 1A–Death Receptor 3 Pathway Promotes Macrophage Foam Cell Formation In Vitro. *The Journal of Immunology,* 184**,** 5827-5834.

MCMAHON, T. P. & ANDERS, D. G. 2002. Interactions between human cytomegalovirus helicase-primase proteins. *Virus Research,* 86**,** 39-52.

MCSHARRY, B., JONES, C., SKINNER, J., KIPLING, D. & WILKINSON, G. 2001. Human telomerase reverse transcriptase-immortalized MRC-5 and HCA2 human fibroblasts are fully permissive for human cytomegalovirus. *Journal of General Virology,* 82**,** 855-863.

MCSHARRY, B. P., TOMASEC, P., NEALE, M. L. & WILKINSON, G. W. G. 2003. The most abundantly transcribed human cytomegalovirus gene (β2. 7) is non-essential for growth in vitro. *Journal of General Virology,* 84**,** 2511-2516.

MEDENBACH, J., SEILER, M. & HENTZE, M. W. 2011. Translational control via protein-regulated upstream open reading frames. *Cell,* 145**,** 902-913.

MEYERS, J. D., FLOURNOY, N. & THOMAS, E. D. 1986. Risk factors for cytomegalovirus infection after human marrow transplantation. *Journal of Infectious Diseases,* 153**,** 478-488.

MICHAELIS, M., DOERR, H. W. & CINATL JR, J. 2009. The story of human cytomegalovirus and cancer: increasing evidence and open questions. *Neoplasia (New York, NY),* 11**,** 1.

MICHEL, D., PAVIĆ, I., ZIMMERMANN, A., HAUPT, E., WUNDERLICH, K., HEUSCHMID, M. & MERTENS, T. 1996. The UL97 gene product of human cytomegalovirus is an early-late protein with a nuclear localization but is not a nucleoside kinase. *Journal of virology,* 70**,** 6340-6346.

MICHELSON, S., DAL MONTE, P., ZIPETO, D., BODAGHI, B., LAURENT, L., OBERLIN, E., ARENZANA-SEISDEDOS, F., VIRELIZIER, J. L. & LANDINI, M. P. 1997. Modulation of RANTES production by human cytomegalovirus infection of fibroblasts. *Journal of virology,* 71**,** 6495-6500.

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. *Journal of general virology,* 88**,** 2642-2650.

MILLER, D. M., CEBULLA, C. M., RAHILL, B. M. & SEDMAK, D. D. Cytomegalovirus and transcriptional down-regulation of major histocompatibility complex class II expression. 2001. Elsevier, 11-18.

MILLER, M. S., FURLONG, W. E., PENNELL, L., GEADAH, M. & HERTEL, L. 2010. RASCAL is a new human cytomegalovirus-encoded protein that localizes to the nuclear lamina and in cytoplasmic vesicles at late times postinfection. *Journal of virology,* 84**,** 6483-6496.

MILNE, R., PATERSON, D. A. & BOOTH, J. C. 1998. Human cytomegalovirus glycoprotein H/glycoprotein L complex modulates fusion-from-without. *Journal of General Virology,* 79**,** 855-865.

MITCHELL, D. 2009. *Human cytomegalovirus UL2829 protein is required for efficient expression of the viral genome*, Princeton University.

MITCHELL, D. P., SAVARYN, J. P., MOORMAN, N. J., SHENK, T. & TERHUNE, S. S. 2009. Human cytomegalovirus UL28 and UL29 open reading frames encode a spliced mRNA and stimulate accumulation of immediate-early RNAs. *Journal of virology,* 83**,** 10187-10197.

MOCARSKI, E. 1993. Cytomegalovirus biology and replication. *The human herpesviruses.*

MOCARSKI, E. S. 2001. *Cytomegaloviruses and their replication*.

MOCARSKI, E. S., KEMBLE, G. W., LYLE, J. M. & GREAVES, R. F. 1996. A deletion mutant in the human cytomegalovirus gene encoding IE1 (491aa) is replication defective due to a failure in autoregulation. *Proceedings of the National Academy of Sciences,* 93**,** 11321.

MOCARSKI, E. S., PEREIRA, L. & MCCORMICK, A. L. 1988. Human cytomegalovirus ICP22, the product of the HWLF1 reading frame, is an early nuclear protein that is released from cells. *The Journal of general virology,* 69**,** 2613.

MOKROS, T., REHM, A., DROESE, J., OPPERMANN, M., LIPP, M. & HÖPKEN, U. E. 2002. Surface expression and endocytosis of the human cytomegalovirus-encoded chemokine receptor US28 is regulated by agonist-independent phosphorylation. *Journal of Biological Chemistry,* 277**,** 45122-45128.

MONTAG, C., WAGNER, J., GRUSKA, I. & HAGEMEIER, C. 2006. Human cytomegalovirus blocks tumor necrosis factor alpha-and interleukin-1β-mediated NF-κB signaling. *Journal of virology,* 80**,** 11686-11698.

MONTAG, C., WAGNER, J. A., GRUSKA, I., VETTER, B., WIEBUSCH, L. & HAGEMEIER, C. 2011. The Latency-Associated UL138 Gene Product of Human Cytomegalovirus Sensitizes Cells to Tumor Necrosis Factor Alpha (TNF-α) Signaling by Upregulating TNF-α Receptor 1 Cell Surface Expression. *Journal of virology,* 85**,** 11409-11421.

MOORMAN, N. J., CRISTEA, I. M., TERHUNE, S. S., ROUT, M. P., CHAIT, B. T. & SHENK, T. 2008. Human cytomegalovirus protein UL38 inhibits host cell stress responses by antagonizing the tuberous sclerosis protein complex. *Cell host & microbe,* 3**,** 253-262.

MORRIS, D. R. & GEBALLE, A. P. 2000. Upstream open reading frames as regulators of mRNA translation. *Molecular and Cellular Biology,* 20**,** 8635-8642.

MULLBERG, J., HSU, M., RAUCH, C., GERHART, M., KAYKAS, A. & COSMAN, D. 1999. The R27080 glycoprotein is abundantly secreted from human cytomegalovirus-infected fibroblasts. *Journal of general virology,* 80**,** 437-440.

MÜLLBERG, J., HSU, M., RAUCH, C., GERHART, M., KAYKAS, A. & COSMAN, D. 1999. The R27080 glycoprotein is abundantly secreted from human cytomegalovirus-infected fibroblasts. *Journal of General Virology,* 80**,** 437-440.

MURAKAMI, P. & MCCAMAN, M. T. 1999. Quantitation of adenovirus DNA and virus particles with the PicoGreen fluorescent Dye. *Analytical biochemistry,* 274**,** 283-288.

MURPHY, E., YU, D., GRIMWOOD, J., SCHMUTZ, J., DICKSON, M., JARVIS, M. A., HAHN, G., NELSON, J. A., MYERS, R. M. & SHENK, T. E. 2003. Coding potential of laboratory and clinical strains of human cytomegalovirus. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 14976.

MURPHY, J. C., FISCHLE, W., VERDIN, E. & SINCLAIR, J. H. 2002. Control of cytomegalovirus lytic gene expression by histone acetylation. *The EMBO journal,* 21**,** 1112-1120.

MYEKU, N. & FIGUEIREDO-PEREIRA, M. E. 2011. Dynamics of the Degradation of Ubiquitinated Proteins by Proteasomes and Autophagy ASSOCIATION WITH SEQUESTOSOME 1/p62. *Journal of Biological Chemistry,* 286**,** 22426-22440.

NELMS, B. L. & LABOSKY, P. A. 2011. A predicted hairpin cluster correlates with barriers to PCR, sequencing and possibly BAC recombineering. *Scientific reports,* 1.

NEVELS, M., BRUNE, W. & SHENK, T. 2004. SUMOylation of the human cytomegalovirus 72-kilodalton IE1 protein facilitates expression of the 86-kilodalton IE2 protein and promotes viral replication. *Journal of virology,* 78**,** 7803-7812.

NGUYEN, N. L., LOVELAND, A. N. & GIBSON, W. 2008. Nuclear localization sequences in cytomegalovirus capsid assembly proteins (UL80 proteins) are required for virus production: inactivating NLS1, NLS2, or both affects replication to strikingly different extents. *Journal of virology,* 82**,** 5381-5389.

NOGALSKI, M. T., PODDUTURI, J. P., DEMERITT, I. B., MILFORD, L. E. & YUROCHKO, A. D. 2007. The human cytomegalovirus virion possesses an activated casein kinase II that allows for the rapid phosphorylation of the inhibitor of NF-κB, IκBα. *Journal of virology,* 81**,** 5305-5314.

O'BRIEN, V. 1998. Viruses and apoptosis. *Journal of General Virology,* 79**,** 1833-1845.

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.

OKUMURA, A., PITHA, P. M., YOSHIMURA, A. & HARTY, R. N. 2010. Interaction between Ebola virus glycoprotein and host toll-like receptor 4 leads to induction of proinflammatory cytokines and SOCS1. *Journal of virology,* 84**,** 27-33.

PARI, G. S., KACICA, M. & ANDERS, D. 1993. Open reading frames UL44, IRS1/TRS1, and UL36-38 are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA synthesis. *Journal of virology,* 67**,** 2575-2582.

PARK, B., OH, H., LEE, S., SONG, Y., SHIN, J., SUNG, Y. C., HWANG, S. Y. & AHN, K. 2002. The MHC class I homolog of human cytomegalovirus is resistant to down-regulation mediated by the unique short region protein (US) 2, US3, US6, and US11 gene products. *The Journal of Immunology,* 168**,** 3464-3469.

PATERSON, D. A., DYER, A. P., MILNE, R. S. B., SEVILLA-REYES, E. & GOMPELS, U. A. 2002. A role for human cytomegalovirus glycoprotein O (gO) in cell fusion and a new hypervariable locus. *Virology,* 293**,** 281-294.

PATRONE, M., PERCIVALLE, E., SECCHI, M., FIORINA, L., PEDRALI-NOY, G., ZOPPÉ, M., BALDANTI, F., HAHN, G., KOSZINOWSKI, U. H. & MILANESI, G. 2003. The human cytomegalovirus UL45 gene product is a late, virion-associated protein and influences virus growth at low multiplicities of infection. *Journal of general virology,* 84**,** 3359-3370.

PATRONE, M., SECCHI, M., FIORINA, L., IERARDI, M., MILANESI, G. & GALLINA, A. 2005. Human cytomegalovirus UL130 protein promotes endothelial cell infection through a producer cell modification of the virion. *Journal of virology,* 79**,** 8361-8373.

PATTERSON, C. E. & SHENK, T. 1999. Human cytomegalovirus UL36 protein is dispensable for viral replication in cultured cells. *Journal of virology,* 73**,** 7126-7131.

PAWELEC, G., AKBAR, A., BEVERLEY, P., CARUSO, C., DERHOVANESSIAN, E., FÜLÖP, T., GRIFFITHS, P., GRUBECK-LOEBENSTEIN, B., HAMPRECHT, K. & JAHN, G. 2010. Immunosenescence and Cytomegalovirus: where do we stand after a decade. *Immun Ageing,* 7**,** 13.

PAYA, C. V., WOLD, A. & SMITH, T. F. 1987. Detection of cytomegalovirus infections in specimens other than urine by the shell vial assay and conventional tube cell cultures. *Journal of clinical microbiology,* 25**,** 755-757.

PENFOLD, M. E. T., DAIRAGHI, D. J., DUKE, G. M., SAEDERUP, N., MOCARSKI, E. S., KEMBLE, G. W. & SCHALL, T. J. 1999. Cytomegalovirus encodes a potent α chemokine. *Proceedings of the National Academy of Sciences,* 96**,** 9839-9844.

PERNG, Y. C., QIAN, Z., FEHR, A. R., XUAN, B. & YU, D. 2011. The human cytomegalovirus gene UL79 is required for the accumulation of late viral transcripts. *Journal of virology,* 85**,** 4841-4852.

PETER, M., HELLBARDT, S., SCHWARTZ-ALBIEZ, R., WESTENDORP, M., WALCZAK, H., MOLDENHAUER, G., GRELL, M. & KRAMMER, P. 1995. Cell surface sialylation plays a

role in modulating sensitivity towards APO-1-mediated apoptotic cell death. *Cell death and differentiation,* 2**,** 163.
PETER, M. E. & KRAMMER, P. 2003. The CD95 (APO-1/Fas) DISC and beyond. *Cell Death & Differentiation,* 10**,** 26-35.
PETRUCELLI, A., RAK, M., GRAINGER, L. & GOODRUM, F. 2009. Characterization of a novel Golgi apparatus-localized latency determinant encoded by human cytomegalovirus. *Journal of virology,* 83**,** 5615-5629.
PETRUCELLI, A., UMASHANKAR, M., ZAGALLO, P., RAK, M. & GOODRUM, F. 2012. Interactions between proteins encoded within the human cytomegalovirus UL133-UL138 locus. *Journal of virology,* 86**,** 8653-8662.
PFEFFER, S., SEWER, A., LAGOS-QUINTANA, M., SHERIDAN, R., SANDER, C., GRÄSSER, F. A., VAN DYK, L. F., HO, C. K., SHUMAN, S. & CHIEN, M. 2005. Identification of microRNAs of the herpesvirus family. *Nature methods,* 2**,** 269-276.
PIZZORNO, M. C., O'HARE, P., SHA, L., LAFEMINA, R. L. & HAYWARD, G. 1988. trans-activation and autoregulation of gene expression by the immediate-early region 2 gene products of human cytomegalovirus. *Journal of virology,* 62**,** 1167-1179.
PLACHTER, B., TRAUPE, B., ALBRECHT, J. & JAHN, G. 1988. Abundant 5 kb RNA of human cytomegalovirus without a major translational reading frame. *The Journal of general virology,* 69**,** 2251.
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. *Journal of virology,* 72**,** 7722-7732.
PLESKOFF, O., TRÉBOUTE, C. & ALIZON, M. 1998. The cytomegalovirus-encoded chemokine receptor US28 can enhance cell-cell fusion mediated by different viral proteins. *Journal of virology,* 72**,** 6389-6397.
POMA, E. E., KOWALIK, T. F., ZHU, L., SINCLAIR, J. H. & HUANG, E. S. 1996. The human cytomegalovirus IE1-72 protein interacts with the cellular p107 protein and relieves p107-mediated transcriptional repression of an E2F-responsive promoter. *Journal of virology,* 70**,** 7867-7877.
POOLE, E., DALLAS, S. R. M. G., COLSTON, J., JOSEPH, R. S. V. & SINCLAIR, J. 2011. Virally induced changes in cellular microRNAs maintain latency of human cytomegalovirus in CD34+ progenitors. *Journal of General Virology,* 92**,** 1539-1549.
POON, A. & ROIZMAN, B. 1993. Characterization of a temperature-sensitive mutant of the UL15 open reading frame of herpes simplex virus 1. *Journal of virology,* 67**,** 4497-4503.
PRICHARD, M., PENFOLD, M., DUKE, G., SPAETE, R. & KEMBLE, G. 2001. A review of genetic differences between limited and extensively passaged human cytomegalovirus strains. *Reviews in medical virology,* 11**,** 191-200.
PRICHARD, M. N., BRITT, W. J., DAILY, S. L., HARTLINE, C. B. & KERN, E. R. 2005a. Human cytomegalovirus UL97 kinase is required for the normal intranuclear distribution of pp65 and virion morphogenesis. *Journal of virology,* 79**,** 15494-15502.
PRICHARD, M. N., LAWLOR, H., DUKE, G. M., MO, C., WANG, Z., DIXON, M., KEMBLE, G. & KERN, E. R. 2005. Human cytomegalovirus uracil DNA glycosylase associates with ppUL44 and accelerates the accumulation of viral DNA. *Virology journal,* 2**,** 55.
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. *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. 2007. The human cytomegalovirus MHC class I homolog UL18 inhibits LIR-1+ but activates LIR-1–NK cells. *The Journal of Immunology,* 178**,** 4473.

QIAN, Z., XUAN, B. & YU, D. 2008. The full-length protein encoded by human cytomegalovirus gene UL117 is required for the proper maturation of viral replication compartments. *Journal of virology,* 82**,** 3452-3465.

QUINNAN JR, G. V., DELERY, M., ROOK, A. H., FREDERICK, W. R., EPSTEIN, J. A. Y. S., MANISCHEWITZ, J. F., JACKSON, L., RAMSEY, K. M., MITTAL, K. & PLOTKIN, S. A. 1984. Comparative virulence and immunogenicity of the Towne strain and a nonattenuated strain of cytomegalovirus. *Annals of internal medicine,* 101**,** 478-483.

RASMUSSEN, L., MATKIN, C., SPAETE, R., PACHL, C. & MERIGAN, T. C. 1991. Antibody response to human cytomegalovirus glycoproteins gB and gH after natural infection in humans. *Journal of Infectious Diseases,* 164**,** 835-842.

RAWLINSON, W. D. & BARRELL, B. G. 1993. Spliced transcripts of human cytomegalovirus. *Journal of virology,* 67**,** 5502-5513.

RE, M. C., LANDINI, M. P., COPPOLECCHIA, P., FURLINI, G. & LA PLACA, M. 1985. A 28 000 Molecular Weight Human Cytomegalovirus Structural Polypeptide Studied by means of a Specific Monoclonal Antibody. *Journal of general virology,* 66**,** 2507-2511.

REAL, F. X., RETTIG, W. J., CHESA, P. G., MELAMED, M. R., OLD, L. J. & MENDELSOHN, J. 1986. Expression of epidermal growth factor receptor in human cultured cells and tissues: relationship to cell lineage and stage of differentiation. *Cancer research,* 46**,** 4726.

REEVES, M., MACARY, P., LEHNER, P., SISSONS, J. & SINCLAIR, J. 2005. 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.

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. *Journal of virology,* 80**,** 9998-10009.

REEVES, M. B., DAVIES, A. A., MCSHARRY, B. P., WILKINSON, G. W. & SINCLAIR, J. H. 2007. Complex I binding by a virally encoded RNA regulates mitochondria-induced cell death. *Science's STKE,* 316**,** 1345.

REEVES, M. B. & SINCLAIR, J. H. 2010. Analysis of latent viral gene expression in natural and experimental latency models of human cytomegalovirus and its correlation with histone modifications at a latent promoter. *Journal of General Virology,* 91**,** 599-604.

REHM, A., ENGELSBERG, A., TORTORELLA, D., KÖRNER, I. J., LEHMANN, I., PLOEGH, H. L. & HÖPKEN, U. E. 2002. Human cytomegalovirus gene products US2 and US11 differ in their ability to attack major histocompatibility class I heavy chains in dendritic cells. *Journal of virology,* 76**,** 5043-5050.

RIBBERT, H. 1904. Ueber protozoenartige Zellen in der Niere eines syphilitischen Neugeborenen und in der Parotis von Kidern. *Zentralbl Allg Pathol,* 15**,** 945-948.

RICE, G., SCHRIER, R. D. & OLDSTONE, M. 1984. Cytomegalovirus infects human lymphocytes and monocytes: virus expression is restricted to immediate-early gene products. *Proceedings of the National Academy of Sciences,* 81**,** 6134.

RIDDELL, S. R., WATANABE, K. S., GOODRICH, J. M., LI, C. R., AGHA, M. E. & GREENBERG, P. D. 1992. Restoration of viral immunity in immunodeficient humans by the adoptive transfer of T cell clones. *Science,* 257**,** 238-241.

ROBY, C. & GIBSON, W. 1986. Characterization of phosphoproteins and protein kinase activity of virions, noninfectious enveloped particles, and dense bodies of human cytomegalovirus. *Journal of virology,* 59**,** 714-727.

ROIZMAN, B. 1980. GENOME VARIATION AND EVOLUTION AMONG HERPES VIRUSES*. *Annals of the New York Academy of Sciences,* 354**,** 472-483.

ROMANOWSKI, M. J., GARRIDO-GUERRERO, E. & SHENK, T. 1997. pIRS1 and pTRS1 are present in human cytomegalovirus virions. *Journal of virology,* 71**,** 5703-5705.

ROMANOWSKI, M. J. & SHENK, T. 1997. Characterization of the human cytomegalovirus irs1 and trs1 genes: a second immediate-early transcription unit within irs1 whose product antagonizes transcriptional activation. *Journal of virology,* 71**,** 1485-1496.

ROWE, W., HARTLEY, J., WATERMAN, S., TURNER, H. & HUEBNER, R. 1956. Cytopathogenic agent resembling human salivary gland virus recovered from tissue cultures of human adenoids. *Experimental Biology and Medicine,* 92**,** 418.

ROWE, W. P., HUEBNER, R. J., GILMORE, L. K., PARROTT, R. H. & WARD, T. G. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. 1953. Royal Society of Medicine, 570-573.

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.

RYCKMAN, B. J., RAINISH, B. L., CHASE, M. C., BORTON, J. A., NELSON, J. A., JARVIS, M. A. & JOHNSON, D. C. 2008. Characterization of the human cytomegalovirus gH/gL/UL128-131 complex that mediates entry into epithelial and endothelial cells. *Journal of virology,* 82**,** 60-70.

SAILESH, H., NELL, L., REZA, Z. M. & GREGORY, S. 2007. Differential inhibition of human cytomegalovirus (HCMV) by toll-like receptor ligands mediated by interferon-beta in human foreskin fibroblasts and cervical tissue. *Virology Journal,* 4.

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.

SAMBUCETTI, L. C., CHERRINGTON, J. M., WILKINSON, G. & MOCARSKI, E. S. 1989. NF-kappa B activation of the cytomegalovirus enhancer is mediated by a viral transactivator and by T cell stimulation. *The EMBO journal,* 8**,** 4251.

SAMPAIO, K. L., CAVIGNAC, Y., STIERHOF, Y. D. & SINZGER, C. 2005. Human cytomegalovirus labeled with green fluorescent protein for live analysis of intracellular particle movements. *Journal of virology,* 79**,** 2754-2767.

SANCHEZ, V., ANGELETTI, P. C., ENGLER, J. & BRITT, W. 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. 2000. Accumulation of virion tegument and envelope proteins in a stable cytoplasmic compartment during human cytomegalovirus replication: characterization of a potential site of virus assembly. *Journal of virology,* 74**,** 975-986.

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. *Journal of virology,* 70**,** 7398-7413.

SAROV, I. & ABADY, I. 1975. The morphogenesis of human cytomegalovirus:: Isolation and polypeptide characterization of cytomegalovirions and dense bodies. *Virology,* 66**,** 464-473.

SCALZO, A. A., FORBES, C. A., SMITH, L. M. & LOH, L. C. 2009. Transcriptional analysis of human cytomegalovirus and rat cytomegalovirus homologues of the M73/M73. 5 spliced gene family. *Archives of virology,* 154**,** 65-75.

SCHAUFLINGER, M., FISCHER, D., SCHREIBER, A., CHEVILLOTTE, M., WALTHER, P., MERTENS, T. & VON EINEM, J. 2011. The tegument protein UL71 of human cytomegalovirus is involved in late envelopment and affects multivesicular bodies. *Journal of virology,* 85**,** 3821-3832.

SCHIRM, J., TIMMERIJE, W., VAN DER BIJ, W., WILTERDINK, J. B., TEGZESS, A. M., VAN SON, W. J. & SCHRODER, F. P. 1987. Rapid detection of infectious cytomegalovirus in blood with the aid of monoclonal antibodies. *Journal of medical virology,* 23**,** 31-40.

SCHLEISS, M. R., DEGNIN, C. & GEBALLE, A. P. 1991. Translational control of human cytomegalovirus gp48 expression. *Journal of virology,* 65**,** 6782-6789.
SCHMOLKE, S., DRESCHER, P., JAHN, G. & PLACHTER, B. 1995a. Nuclear targeting of the tegument protein pp65 (UL83) of human cytomegalovirus: an unusual bipartite nuclear localization signal functions with other portions of the protein to mediate its efficient nuclear transport. *Journal of virology,* 69**,** 1071-1078.
SCHMOLKE, S., KERN, H. F., DRESCHER, P., JAHN, G. & PLACHTER, B. 1995b. The dominant phosphoprotein pp65 (UL83) of human cytomegalovirus is dispensable for growth in cell culture. *Journal of virology,* 69**,** 5959-5968.
SCHOPPEL, K., SCHMIDT, C., EINSELE, H., HEBART, H. & MACH, M. 1998. Kinetics of the antibody response against human cytomegalovirus-specific proteins in allogeneic bone marrow transplant recipients. *Journal of Infectious Diseases,* 178**,** 1233-1243.
SCULLY, A. L., SOMMER, M. H., SCHWARTZ, R. & SPECTOR, D. H. 1995. The human cytomegalovirus IE2 86-kilodalton protein interacts with an early gene promoter via site-specific DNA binding and protein-protein associations. *Journal of virology,* 69**,** 6533-6540.
SEIBENHENER, M. L., BABU, J. R., GEETHA, T., WONG, H. C., KRISHNA, N. R. & WOOTEN, M. W. 2004. Sequestosome 1/p62 is a polyubiquitin chain binding protein involved in ubiquitin proteasome degradation. *Molecular and Cellular Biology,* 24**,** 8055-8068.
SEVERI, B., LANDINI, M. & GOVONI, E. 1988. Human cytomegalovirus morphogenesis: an ultrastructural study of the late cytoplasmic phases. *Archives of virology,* 98**,** 51-64.
SHATNYEVA, O. M., KUBARENKO, A. V., WEBER, C. E. M., PAPPA, A., SCHWARTZ-ALBIEZ, R., WEBER, A. N. R., KRAMMER, P. H. & LAVRIK, I. N. 2011. Modulation of the CD95-induced apoptosis: the role of CD95 N-glycosylation. *PloS one,* 6**,** e19927.
SHELBOURN, S., KOTHARI, S., SISSONS, J. & SINCLAIR, J. 1989. Repression of human cytomegalovirus gene expression asscoiated with a novel immediate early regulatory region binding factor. *Nucleic acids research,* 17**,** 9165-9171.
SHEN, A., LEI, J., YANG, E., PEI, Y., CHEN, Y. C., GONG, H., XIAO, G. & LIU, F. 2011. Human cytomegalovirus primase UL70 specifically interacts with cellular factor Snapin. *Journal of virology,* 85**,** 11732-11741.
SHIKHAGAIE, M., MERCÉ-MALDONADO, E., ISERN, E., MUNTASELL, A., ALBÀ, M. M., LÓPEZ-BOTET, M., HENGEL, H. & ANGULO, A. 2012a. The HCMV-specific UL1 gene encodes a late phase glycoprotein incorporated in the virion envelope. *Journal of virology*.
SHIKHAGAIE, M., MERCÉ-MALDONADO, E., ISERN, E., MUNTASELL, A., ALBÀ, M. M., LÓPEZ-BOTET, M., HENGEL, H. & ANGULO, A. 2012b. The Human Cytomegalovirus-Specific UL1 Gene Encodes a Late-Phase Glycoprotein Incorporated in the Virion Envelope. *Journal of virology,* 86**,** 4091-4101.
SHIN, J., PARK, B., LEE, S., KIM, Y., BIEGALKE, B. J., KANG, S. & AHN, K. 2006. A short isoform of human cytomegalovirus US3 functions as a dominant negative inhibitor of the full-length form. *Journal of virology,* 80**,** 5397-5404.
SHIRAKATA, M., TERAUCHI, M., ABLIKIM, M., IMADOME, K. I., HIRAI, K., ASO, T. & YAMANASHI, Y. 2002. Novel immediate-early protein IE19 of human cytomegalovirus activates the origin recognition complex I promoter in a cooperative manner with IE72. *Journal of virology,* 76**,** 3158-3167.
SHISLER, J., YANG, C., WALTER, B., WARE, C. F. & GOODING, L. R. 1997. The adenovirus E3-10.4 K/14.5 K complex mediates loss of cell surface Fas (CD95) and resistance to Fas-induced apoptosis. *Journal of virology,* 71**,** 8299-8306.
SHUNG, C. Y. & SUNTER, G. 2009. Regulation of Tomato golden mosaic virus AL2 and AL3 gene expression by a conserved upstream open reading frame. *Virology,* 383**,** 310-318.
SILVA, M. C., YU, Q. C., ENQUIST, L. & SHENK, T. 2003. Human cytomegalovirus UL99-encoded pp28 is required for the cytoplasmic envelopment of tegument-associated capsids. *Journal of virology,* 77**,** 10594-10605.

SINCLAIR, J. 2008. Human cytomegalovirus: latency and reactivation in the myeloid lineage. *Journal of clinical virology,* 41**,** 180-185.
SINCLAIR, J. 2010a. Chromatin structure regulates human cytomegalovirus gene expression during latency, reactivation and lytic infection. *Biochimica et biophysica acta,* 1799**,** 286.
SINCLAIR, J. 2010b. Chromatin structure regulates human cytomegalovirus gene expression during latency, reactivation and lytic infection. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms,* 1799**,** 286-295.
SINCLAIR, J. & SISSONS, P. 2006. Latency and reactivation of human cytomegalovirus. *Journal of General Virology,* 87**,** 1763-1779.
SINZGER, C., GREFTE, A., PLACHTER, B., GOUW, A. S. 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. *Journal of General Virology,* 76**,** 741.
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. *Journal of General Virology,* 89**,** 359-368.
SINZGER, C., KAHL, M., LAIB, K., KLINGEL, K., RIEGER, P., PLACHTER, B. & JAHN, G. 2000. Tropism of human cytomegalovirus for endothelial cells is determined by a post-entry step dependent on efficient translocation to the nucleus. *Journal of General Virology,* 81**,** 3021-3035.
SINZGER, C., PLACHTER, B., GREFTE, A. & JAHN, G. 1996. Tissue macrophages are infected by human cytomegalovirus in vivo. *Journal of Infectious Diseases,* 173**,** 240.
SINZGER, C., SCHMIDT, K., KNAPP, J., KAHL, M., BECK, R., WALDMAN, J., HEBART, H., EINSELE, H. & JAHN, G. 1999. Modification of human cytomegalovirus tropism through propagation in vitro is associated with changes in the viral genome. *Journal of General Virology,* 80**,** 2867-2877.
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. *Proceedings of the National Academy of Sciences,* 98**,** 7829.
SLOBEDMAN, B. & MOCARSKI, E. S. 1999. Quantitative analysis of latent human cytomegalovirus. *Journal of virology,* 73**,** 4806-4812.
SMITH, A. & WEIDMAN, F. 1910. Infection of a Still-born Infant by an Amebiform Protozoan. *Univ. of Penn. Med. Bull.,* 23.
SMITH, A. & WEIDMAN, F. 1914. Further Note upon the Occurrence of Entamoeba Mortinatalium as a Human Parasite. *Amer. J. Trop. Dis. and Prev. Med,* 2**,** 256-259.
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. *Archives of ophthalmology,* 116**,** 178.
SMITH, J. A., JAIRATH, S., CRUTE, J. J. & PARI, G. S. 1996. Characterization of the human cytomegalovirus UL105 gene and identification of the putative helicase protein. *Virology,* 220**,** 251-255.
SMITH, J. A. & PARI, G. S. 1995. Expression of human cytomegalovirus UL36 and UL37 genes is required for viral DNA replication. *Journal of virology,* 69**,** 1925-1931.
SMITH, K. O. & RASMUSSEN, L. 1963. Morphology of cytomegalovirus (salivary gland virus). *Journal of bacteriology,* 85**,** 1319-1325.
SMITH, M. G. Propagation in tissue cultures of a cytopathogenic virus from human salivary gland virus (SGV) disease. 1956. Royal Society of Medicine, 424-430.
SMITH, M. G. 1959. The salivary gland viruses of man and animals (cytomegalic inclusion disease). *Prog. Med. Virol,* 2**,** 171-202.

SNYDMAN, D. R., WERNER, B. G., HEINZE-LACEY, B., BERARDI, V. P., TILNEY, N. L., KIRKMAN, R. L., MILFORD, E. L., CHO, S. I., BUSH JR, H. L. & LEVEY, A. S. 1987. Use of cytomegalovirus immune globulin to prevent cytomegalovirus disease in renal-transplant recipients. *New England Journal of Medicine,* 317**,** 1049-1054.

SÖDERBERG-NAUCLÉR, C. 2008. HCMV microinfections in inflammatory diseases and cancer. *Journal of clinical virology,* 41**,** 218-223.

SOMMER, M. H., SCULLY, A. L. & SPECTOR, D. H. 1994. Transactivation by the human cytomegalovirus IE2 86-kilodalton protein requires a domain that binds to both the TATA box-binding protein and the retinoblastoma protein. *Journal of virology,* 68**,** 6223-6231.

SOROCEANU, L., AKHAVAN, A. & COBBS, C. S. 2008. Platelet-derived growth factor-&agr; receptor activation is required for human cytomegalovirus infection. *Nature,* 455**,** 391-395.

SOURVINOS, G., TAVALAI, N., BERNDT, A., SPANDIDOS, D. A. & STAMMINGER, T. 2007. Recruitment of human cytomegalovirus immediate-early 2 protein onto parental viral genomes in association with ND10 in live-infected cells. *Journal of virology,* 81**,** 10123-10136.

SPADERNA, S., BLESSING, H., BOGNER, E., BRITT, W. & MACH, M. 2002. Identification of glycoprotein gpTRL10 as a structural component of human cytomegalovirus. *Journal of virology,* 76**,** 1450-1460.

SPADERNA, S., KROPFF, B., KÖDEL, Y., SHEN, S., COLEY, S., LU, S., BRITT, W. & MACH, M. 2005. Deletion of gpUL132, a structural component of human cytomegalovirus, results in impaired virus replication in fibroblasts. *Journal of virology,* 79**,** 11837-11847.

SPAETE, R. R., GEHRZ, R. C. & LANDINI, M. P. 1994. Human cytomegalovirus structural proteins. *Journal of General Virology,* 75**,** 3287-3308.

SPAETE, R. R. & MOCARSKI, E. S. 1985. The alpha sequence of the cytomegalovirus genome functions as a cleavage/packaging signal for herpes simplex virus defective genomes. *Journal of virology,* 54**,** 817.

SPAETE, R. R. & MOCARSKI, E. S. 1987. Insertion and deletion mutagenesis of the human cytomegalovirus genome. *Proceedings of the National Academy of Sciences,* 84**,** 7213.

SPAETE, R. R., PEROT, K., SCOTT, P. I., NELSON, J. A., STINSKI, M. F. & PACHL, C. 1993. Coexpression of truncated human cytomegalovirus gH with the UL115 gene product or the truncated human fibroblast growth factor receptor results in transport of gH to the cell surface. *Virology,* 193**,** 853-861.

SPECTOR, D. H. 1996. Activation and regulation of human cytomegalovirus early genes. *Intervirology,* 39**,** 361-377.

SPEIR, E., MODALI, R., HUANG, E. S., LEON, M. B., SHAWL, F., FINKEL, T. & EPSTEIN, S. E. 1994. Potential role of human cytomegalovirus and p53 interaction in coronary restenosis. *Science,* 265**,** 391-394.

STAGNO, S., PASS, R., DWORSKY, M. & ALFORD, C. Congenital and perinatal cytomegalovirus infections. 1983. 31.

STAGNO, S., PASS, R., REYNOLDS, D., MOORE, M., NAHMIAS, A. & ALFORD, C. 1980a. Comparative study of diagnostic procedures for congenital cytomegalovirus infection. *Pediatrics,* 65**,** 251-257.

STAGNO, S., REYNOLDS, D. W., PASS, R. F. & ALFORD, C. A. 1980b. Breast milk and the risk of cytomegalovirus infection. *New England Journal of Medicine,* 302**,** 1073-1076.

STAMMINGER, T., GSTAIGER, M., WEINZIERL, K., LORZ, K., WINKLER, M. & SCHAFFNER, W. 2002. Open reading frame UL26 of human cytomegalovirus encodes a novel tegument protein that contains a strong transcriptional activation domain. *Journal of virology,* 76**,** 4836-4847.

STANNARD, L. M. & HARDIE, D. R. 1991. An Fc receptor for human immunoglobulin G is located within the tegument of human cytomegalovirus. *Journal of virology,* 65**,** 3411-3415.

STANTON, R., WESTMORELAND, D., FOX, J. D., DAVISON, A. J. & WILKINSON, G. W. G. 2005. Stability of human cytomegalovirus genotypes in persistently infected renal transplant recipients. *Journal of medical virology,* 75**,** 42-46.

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. 2008a. Re-engineering adenovirus vector systems to enable high-throughput analyses of gene function. *BioTechniques,* 45**,** 659-62, 664-8.

STANTON, R. J., MCSHARRY, B. P., ARMSTRONG, M., TOMASEC, P. & WILKINSON, G. W. G. 2008b. Re-engineering adenovirus vector systems to enable high-throughput analyses of gene function. *BioTechniques,* 45**,** 659-662.

STENBERG, R., FORTNEY, J., BARLOW, S., MAGRANE, B., NELSON, J. & GHAZAL, P. 1990. Promoter-specific trans activation and repression by human cytomegalovirus immediate-early proteins involves common and unique protein domains. *Journal of virology,* 64**,** 1556-1565.

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. 2007a. Host immune system gene targeting by a viral miRNA. *Science's STKE,* 317**,** 376.

STERN-GINOSSAR, N., ELEFANT, N., ZIMMERMANN, A., WOLF, D. G., SALEH, N., BITON, M., HORWITZ, E., PROKOCIMER, Z., PRICHARD, M. & HAHN, G. 2007b. Host immune system gene targeting by a viral miRNA. *Science,* 317**,** 376.

STERN-GINOSSAR, N., SALEH, N., GOLDBERG, M. D., PRICHARD, M., WOLF, D. G. & MANDELBOIM, O. 2009. Analysis of human cytomegalovirus-encoded microRNA activity during infection. *Journal of virology,* 83**,** 10684-10693.

STERN, H. & FRIEDMANN, I. 1960. Intranuclear formation of cytomegalic inclusion disease virus.

STEVEN, A. & SPEAR, P. 1997. Herpesvirus capsid assembly and envelopment. *Structural biology of viruses. Oxford University Press, New York, NY***,** 312-351.

STINSKI, M. 1983. Molecular biology of cytomegaloviruses. *The herpesviruses,* 2**,** 67-113.

STINSKI, M. & MEIER, J. 2007. Immediate–early viral gene regulation and function. *Human Herpesviruses*

Biology, Therapy, and Immunoprophylaxis. Cambridge: Cambridge University Press.

STINSKI, M. F. 1977. Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. *Journal of virology,* 23**,** 751-767.

STIRK, P. & GRIFFITHS, P. 1987. Use of monoclonal antibodies for the diagnosis of cytomegalovirus infection by the detection of early antigen fluorescent foci (DEAFF) in cell culture. *Journal of medical virology,* 21**,** 329-337.

STRANG, B. L., BOULANT, S., CHANG, L., KNIPE, D. M., KIRCHHAUSEN, T. & COEN, D. M. 2012. Human cytomegalovirus UL44 concentrates at the periphery of replication compartments, the site of viral DNA synthesis. *Journal of virology,* 86**,** 2089-2095.

SU, Y., TESTAVERDE, J. R., DAVIS, C. N., HAYAJNEH, W. A., ADAIR, R. & COLBERG-POLEY, A. M. 2003. Human cytomegalovirus UL37 immediate early target minigene RNAs are accurately spliced and polyadenylated. *Journal of general virology,* 84**,** 29-39.

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. 2005. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. *The Journal of experimental medicine,* 202**,** 673-685.

TAKEUCHI, O. & AKIRA, S. 2009. Innate immunity to virus infection. *Immunological reviews,* 227**,** 75-86.

TALBOT, P. & ALMEIDA, J. D. 1977. Human cytomegalovirus: purification of enveloped virions and dense bodies. *Journal of General Virology,* 36**,** 345.

TAMASHIRO, J. C. & SPECTOR, D. H. 1986. Terminal structure and heterogeneity in human cytomegalovirus strain AD169. *Journal of virology,* 59**,** 591.

TANAKA, K., ZOU, J. P., TAKEDA, K., FERRANS, V. J., SANDFORD, G. R., JOHNSON, T. M., FINKEL, T. & EPSTEIN, S. E. 1999. Effects of human cytomegalovirus immediate-early proteins on p53-mediated apoptosis in coronary artery smooth muscle cells. *Circulation,* 99**,** 1656-1659.

TANDON, R. & MOCARSKI, E. S. 2008. Control of cytoplasmic maturation events by cytomegalovirus tegument protein pp150. *Journal of virology,* 82**,** 9433-9444.

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. *Journal of virology,* 80**,** 8006-8018.

TAYLOR-WIEDEMAN, J., SISSONS, J., BORYSIEWICZ, L. K. & SINCLAIR, J. 1991. Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells. *The Journal of general virology,* 72**,** 2059.

TAYLOR-WIEDEMAN, J., SISSONS, P. & SINCLAIR, J. 1994. Induction of endogenous human cytomegalovirus gene expression after differentiation of monocytes from healthy carriers. *Journal of virology,* 68**,** 1597-1604.

TENNEY, D. J., SANTOMENNA, L. D., GOUDIE, K. B. & COLBERG-POLEY, A. M. 1993. The human cytomegalovirus US3 immediate-early protein lacking the putative transmembrane domain regulates gene expression. *Nucleic acids research,* 21**,** 2931-2937.

TEPPER, C. G., SELDIN, M. F. & MUDRYJ, M. 2000. Fas-mediated apoptosis of proliferating, transiently growth-arrested, and senescent normal human fibroblasts. *Experimental cell research,* 260**,** 9-19.

TERAUCHI, M., KOI, H., HAYANO, C., TOYAMA-SORIMACHI, N., KARASUYAMA, H., YAMANASHI, Y., ASO, T. & SHIRAKATA, M. 2003. Placental extravillous cytotrophoblasts persistently express class I major histocompatibility complex molecules after human cytomegalovirus infection. *Journal of virology,* 77**,** 8187-8195.

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., SCHRÖER, J. & SHENK, T. 2004. RNAs are packaged into human cytomegalovirus virions in proportion to their intracellular concentration. *Journal of virology,* 78**,** 10390-10398.

THEILER, R. N. & COMPTON, T. 2002. Distinct glycoprotein O complexes arise in a post-Golgi compartment of cytomegalovirus-infected cells. *Journal of virology,* 76**,** 2890-2898.

THOMA, C., BORST, E., MESSERLE, M., RIEGER, M., HWANG, J. S. & BOGNER, E. 2006. Identification of the interaction domain of the small terminase subunit pUL89 with the large subunit pUL56 of human cytomegalovirus. *Biochemistry,* 45**,** 8855-8863.

THOMSEN, D. R., STENBERG, R. M., GOINS, W. F. & STINSKI, M. F. 1984. Promoter-regulatory region of the major immediate early gene of human cytomegalovirus. *Proceedings of the National Academy of Sciences,* 81**,** 659.

TOLLEFSON, A. E., HERMISTON, T. W., LICHTENSTEIN, D. L., COLLE, C. F., TRIPP, R. A., DIMITROV, T., TOTH, K., WELLS, C. E., DOHERTY, P. C. & WOLD, W. S. M. 1998. Forced degradation of Fas inhibits apoptosis in adenovirus-infected cells. *Nature,* 392**,** 726-730.

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.

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. 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. *Nature medicine,* 5**,** 1039-1043.

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*.

VAN DE BERG, P. J. E. J., VAN STIJN, A., TEN BERGE, I. J. M. & VAN LIER, R. A. W. 2008. A fingerprint left by cytomegalovirus infection in the human T cell compartment. *Journal of clinical virology,* 41**,** 213-217.

VAN DER BIJ, W., SCHIRM, J., TORENSMA, R., VAN SON, W. & TEGZESS, A. 1988a. Comparison between viremia and antigenemia for detection of cytomegalovirus in blood. *Journal of clinical microbiology,* 26**,** 2531-2535.

VAN DER BIJ, W., TORENSMA, R., VAN SON, W. J., ANEMA, J., SCHIRM, J. & TEGZESS, A. M. 1988b. Rapid immunodiagnosis of active cytomegalovirus infection by monoclonal antibody staining of blood leucocytes. *Journal of medical virology,* 25**,** 179-188.

VANARSDALL, A. L. & JOHNSON, D. C. 2012. Human cytomegalovirus entry into cells. *Current Opinion in Virology*.

VARNUM, S. M., STREBLOW, D. N., MONROE, M. E., SMITH, P., AUBERRY, K. J., PASA-TOLIC, L., WANG, D., CAMP, D. G., RODLAND, K. & WILEY, S. 2004. Identification of proteins in human cytomegalovirus (HCMV) particles: the HCMV proteome. *Journal of virology,* 78**,** 10960.

VERSTREPEN, L., VERHELST, K., CARPENTIER, I. & BEYAERT, R. 2011. TAX1BP1, a ubiquitin-binding adaptor protein in innate immunity and beyond. *Trends in biochemical sciences,* 36**,** 347-354.

VIVIER, E., RAULET, D. H., MORETTA, A., CALIGIURI, M. A., ZITVOGEL, L., LANIER, L. L., YOKOYAMA, W. M. & UGOLINI, S. 2011. Innate or adaptive immunity? The example of natural killer cells. *Science,* 331**,** 44-49.

WAGNER, S., ARNOLD, F., WU, Z., SCHUBERT, A., WALLISER, C., TADAGAKI, K., JOCKERS, R., MERTENS, T. & MICHEL, D. 2012. The 7-transmembrane protein homologue UL78 of the human cytomegalovirus forms oligomers and traffics between the plasma membrane and different intracellular compartments. *Archives of virology***,** 1-15.

WALDMAN, W. J., ROBERTS, W., DAVIS, D., WILLIAMS, M., SEDMAK, D. & STEPHENS, R. 1991. Preservation of natural endothelial cytopathogenicity of cytomegalovirus by propagation in endothelial cells. *Archives of virology,* 117**,** 143-164.

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. 2005. 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. Y., MCSHARRY, B., RETIERE, C., TOMASEC, P., WILLIAMS, S., BORYSIEWICZ, L. K., BRAUD, V. M. & WILKINSON, G. W. G. 2002. UL40-mediated NK evasion during productive infection with human cytomegalovirus. *Proceedings of the National Academy of Sciences,* 99**,** 7570.

WANG, S. K., DUH, C. Y. & CHANG, T. T. 2000. Cloning and identification of regulatory gene UL76 of human cytomegalovirus. *Journal of general virology,* 81**,** 2407-2416.

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**,** 456-461.

WANG, Z. G., RUGGERO, D., RONCHETTI, S., ZHONG, S., GABOLI, M., RIVI, R. & PANDOLFI, P. P. 1998. PML is essential for multiple apoptotic pathways. *Nature genetics,* 20**,** 266-272.
WARMING, S., COSTANTINO, N., JENKINS, N. A. & COPELAND, N. G. 2005. Simple and highly efficient BAC recombineering using galK selection. *Nucleic acids research,* 33**,** e36-e36.
WATHEN, M. W. & STINSKI, M. F. 1982. Temporal patterns of human cytomegalovirus transcription: mapping the viral RNAs synthesized at immediate early, early, and late times after infection. *Journal of virology,* 41**,** 462-477.
WEBER, B., BARGANE FALL, E. H. M., BERGER, A. & DOERR, H. W. 1999. Screening of blood donors for human cytomegalovirus (HCMV) IgG antibody with an enzyme immunoassay using recombinant antigens. *Journal of clinical virology,* 14**,** 173-181.
WEI, X., DECKER, J. M., WANG, S., HUI, H., KAPPES, J. C., WU, X., SALAZAR-GONZALEZ, J. F., SALAZAR, M. G., KILBY, J. M. & SAAG, M. S. 2003. Antibody neutralization and escape by HIV-1. *Nature,* 422**,** 307-312.
WEILAND, K. L., OIEN, N. L., HOMA, F. & WATHEN, M. W. 1994. Functional analysis of human cytomegalovirus polymerase accessory protein. *Virus research,* 34**,** 191-206.
WELLER, T. H. 1971. The cytomegaloviruses: ubiquitous agents with protean clinical manifestations. *New England Journal of Medicine,* 285**,** 203-214.
WELLER, T. H., HANSHAW, J. B. & SCOTT, D. 1960. Serologic Differentiation of Viruses responsible for Cytomegalic Inclusion Disease. *Virology,* 12**,** 130-32.
WELLER, T. H., HANSHAW, J. B. & SCOTT, D. M. E. 1962. Virologic and clinical observations on cytomegalic inclusion disease. *New England Journal of Medicine,* 266**,** 1233-1244.
WELLER, T. H., MACAULEY, J., CRAIG, J. & WIRTH, P. Isolation of intranuclear inclusion producing agents from infants with illnesses resembling cytomegalic inclusion disease. 1957. Royal Society of Medicine, 4-12.
WELLS, D. E., VUGLER, L. G. & BRITT, W. J. 1990. Structural and immunological characterization of human cytomegalovirus gp55-116 (gB) expressed in insect cells. *Journal of general virology,* 71**,** 873-880.
WELTE, S. A., SINZGER, C., LUTZ, S. Z., SINGH-JASUJA, H., SAMPAIO, K. L., EKNIGK, U., RAMMENSEE, H. G. & STEINLE, A. 2002. Selective intracellular retention of virally induced NKG2D ligands by the human cytomegalovirus UL16 glycoprotein. *European journal of immunology,* 33**,** 194-203.
WESTON, K. 1988. An enhancer element in the short unique region of human cytomegalovirus regulates the production of a group of abundant immediate early transcripts. *Virology,* 162**,** 406-416.
WESTSTRATE, M., GEELEN, J. & VAN DER NOORDAA, J. 1980. Human cytomegalovirus DNA: physical maps for the restriction endonucleases BglII, HindIII and XbaI. *Journal of General Virology,* 49**,** 1-21.
WETHMAR, K., SMINK, J. J. & LEUTZ, A. 2010. Upstream open reading frames: molecular switches in (patho) physiology. *Bioessays,* 32**,** 885-893.
WHITE, C. A., STOW, N. D., PATEL, A. H., HUGHES, M. & PRESTON, V. G. 2003. Herpes simplex virus type 1 portal protein UL6 interacts with the putative terminase subunits UL15 and UL28. *Journal of virology,* 77**,** 6351-6358.
WHITE, E. & SPECTOR, D. 2007. Early viral gene expression and function. *Human herpesviruses: biology, therapy, and immunoprophylaxis. Cambridge University Press, Cambridge, United Kingdom.*
WHITE, E. A., CLARK, C. L., SANCHEZ, V. & SPECTOR, D. H. 2004. Small internal deletions in the human cytomegalovirus IE2 gene result in nonviable recombinant viruses with differential defects in viral gene expression. *Journal of virology,* 78**,** 1817-1830.
WIERTZ, E., JONES, T. R., SUN, L., BOGYO, M., GEUZE, H. J. & PLOEGH, H. L. 1996a. The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol. *Cell,* 84**,** 769-780.

WIERTZ, E., TORTORELLA, D., BOGYO, M., YU, J., MOTHES, W., JONES, T. R., RAPOPORT, T. A. & PLOEGH, H. L. 1996b. Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. *Nature,* 384**,** 432-438.

WILKINSON, G., AKRIGG, A. & GREENAWAY, P. 1984. Transcription of the immediate early genes of human cytomegalovirus strain AD169. *Virus Research,* 1**,** 101-116.

WILKINSON, G., 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. 1992. Constitutive and enhanced expression from the CMV major IE promoter in a defective adenovirus vector. *Nucleic acids research,* 20**,** 2233-2239.

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. 2008. Modulation of natural killer cells by human cytomegalovirus. *Journal of clinical virology,* 41**,** 206-212.

WILLCOX, B. E., THOMAS, L. M. & BJORKMAN, P. J. 2003. Crystal structure of HLA-A2 bound to LIR-1, a host and viral major histocompatibility complex receptor. *Nature immunology,* 4**,** 913-919.

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. *Journal of virology,* 84**,** 2585-2596.

WILLS, M. R., ASHIRU, O., REEVES, M. B., OKECHA, G., TROWSDALE, J., TOMASEC, P., WILKINSON, G. W. G., SINCLAIR, J. & SISSONS, J. 2005a. Human cytomegalovirus encodes an MHC class I-like molecule (UL142) that functions to inhibit NK cell lysis. *The Journal of Immunology,* 175**,** 7457.

WILLS, M. R., ASHIRU, O., REEVES, M. B., OKECHA, G., TROWSDALE, J., TOMASEC, P., WILKINSON, G. W. G., SINCLAIR, J. & SISSONS, J. G. P. 2005b. Human cytomegalovirus encodes an MHC class I-like molecule (UL142) that functions to inhibit NK cell lysis. *The Journal of Immunology,* 175**,** 7457-7465.

WILLS, M. R., CARMICHAEL, A. J., MYNARD, K., JIN, X., WEEKES, M. P., PLACHTER, B. & SISSONS, J. 1996. The human cytotoxic T-lymphocyte (CTL) response to cytomegalovirus is dominated by structural protein pp65: frequency, specificity, and T-cell receptor usage of pp65-specific CTL. *Journal of virology,* 70**,** 7569-7579.

WILLS, M. R., CARMICHAEL, A. J., WEEKES, M. P., MYNARD, K., OKECHA, G., HICKS, R. & SISSONS, J. 1999. Human virus-specific CD8+ CTL clones revert from CD45ROhigh to CD45RAhigh in vivo: CD45RAhighCD8+ T cells comprise both naive and memory cells. *The Journal of Immunology,* 162**,** 7080.

WING, B. A., LEE, G. & HUANG, E. S. 1996. The human cytomegalovirus UL94 open reading frame encodes a conserved herpesvirus capsid/tegument-associated virion protein that is expressed with true late kinetics. *Journal of virology,* 70**,** 3339-3345.

WINKLER, M., RICE, S. A. & STAMMINGER, T. 1994. UL69 of human cytomegalovirus, an open reading frame with homology to ICP27 of herpes simplex virus, encodes a transactivator of gene expression. *Journal of virology,* 68**,** 3943-3954.

WINKLER, M. & STAMMINGER, T. 1996. A specific subform of the human cytomegalovirus transactivator protein pUL69 is contained within the tegument of virus particles. *Journal of virology,* 70**,** 8984-8987.

WOLF, D. G., COURCELLE, C. T., PRICHARD, M. N. & MOCARSKI, E. S. 2001. Distinct and separate roles for herpesvirus-conserved UL97 kinase in cytomegalovirus DNA synthesis and encapsidation. *Proceedings of the National Academy of Sciences,* 98**,** 1895-1900.

WOMACK, A. & SHENK, T. 2010. Human cytomegalovirus tegument protein pUL71 is required for efficient virion egress. *MBio,* 1.

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. *Journal of virology,* 71**,** 179-190.

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. *The Journal of biological chemistry,* 281**,** 37652-60.

WRIGHT, D. A. & SPECTOR, D. H. 1989. Posttranscriptional regulation of a class of human cytomegalovirus phosphoproteins encoded by an early transcription unit. *Journal of virology,* 63**,** 3117-3127.

WRIGHT, J. F., KUROSKY, A., PRYZDIAL, E. & WASI, S. 1995. Host cellular annexin II is associated with cytomegalovirus particles isolated from cultured human fibroblasts. *Journal of virology,* 69**,** 4784-4791.

WRIGHT JR, H. T., GOODHEART, C. R. & LIELAUSIS, A. 1964. Human cytomegalovirus. Morphology by negative staining. *Virology,* 23**,** 419-424.

WYATT, J. P., SAXTON, J., LEE, R. & PINKERTON, H. 1950. Generalized cytomegalic inclusion disease. *Journal of Pediatrics,* 36**,** 271-94.

YAMAMOTO, T., SUZUKI, S., RADSAK, K. & HIRAI, K. 1998. The UL112/113 gene products of human cytomegalovirus which colocalize with viral DNA in infected cell nuclei are related to efficient viral DNA replication. *Virus research,* 56**,** 107-14.

YAMASHITA, Y., SHIMOKATA, K., MIZUNO, S., YAMAGUCHI, H. & NISHIYAMA, Y. 1993. Down-regulation of the surface expression of class I MHC antigens by human cytomegalovirus. *Virology,* 193**,** 727-736.

YEAGER, A. S., CARL GRUMET, F., HAFLEIGH, E. B., ARVIN, A. M., BRADLEY, J. S. & PROBER, C. G. 1981. Prevention of transfusion-acquired cytomegalovirus infections in newborn infants. *The Journal of Pediatrics,* 98**,** 281-287.

YEW, K. H., CARPENTER, C., DUNCAN, R. S. & HARRISON, C. J. 2012. Human Cytomegalovirus Induces TLR4 Signaling Components in Monocytes Altering TIRAP, TRAM and Downstream Interferon-Beta and TNF-Alpha Expression. *PloS one,* 7**,** e44500.

YU, D., ELLIS, H. M., LEE, E. C., JENKINS, N. A. & COPELAND, N. G. 2000. An efficient recombination system for chromosome engineering in Escherichia coli. *Proceedings of the National Academy of Sciences,* 97**,** 5978-5983.

YU, X. K., O'CONNOR, C. M., ATANASOV, I., DAMANIA, B., KEDES, D. H. & ZHOU, Z. H. 2003. Three-dimensional structures of the A, B, and C capsids of rhesus monkey rhadinovirus: insights into gammaherpesvirus capsid assembly, maturation, and DNA packaging. *Journal of virology,* 77**,** 13182-13193.

YUROCHKO, A. D., HWANG, E., RASMUSSEN, L., KEAY, S., PEREIRA, L. & HUANG, E. 1997. The human cytomegalovirus UL55 (gB) and UL75 (gH) glycoprotein ligands initiate the rapid activation of Sp1 and NF-kappaB during infection. *Journal of virology,* 71**,** 5051-5059.

ZAIA, J., GALLEZ-HAWKINS, G., CHURCHILL, M., MORTON-BLACKSHERE, A., PANDE, H., ADLER, S., SCHMIDT, G. & FORMAN, S. 1990. Comparative analysis of human cytomegalovirus a-sequence in multiple clinical isolates by using polymerase chain reaction and restriction fragment length polymorphism assays. *Journal of clinical microbiology,* 28**,** 2602.

ZHANG, G., RAGHAVAN, B., KOTUR, M., CHEATHAM, J., SEDMAK, D., COOK, C., WALDMAN, J. & TRGOVCICH, J. 2007. Antisense transcription in the human cytomegalovirus transcriptome. *Journal of virology,* 81**,** 11267.

ZHOU, Z. H., CHEN, D. H., JAKANA, J., RIXON, F. J. & CHIU, W. 1999. Visualization of tegument-capsid interactions and DNA in intact herpes simplex virus type 1 virions. *Journal of virology,* 73**,** 3210-3218.

ZHU, H., CONG, J. P., MAMTORA, G., GINGERAS, T. & SHENK, T. 1998. Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays. *Proceedings of the National Academy of Sciences,* 95**,** 14470.

ZHU, H., CONG, J. P. & SHENK, T. 1997. Use of differential display analysis to assess the effect of human cytomegalovirus infection on the accumulation of cellular RNAs: induction of interferon-responsive RNAs. *Proceedings of the National Academy of Sciences,* 94**,** 13985.

ZHU, H., SHEN, Y. & SHENK, T. 1995. Human cytomegalovirus IE1 and IE2 proteins block apoptosis. *Journal of virology,* 69**,** 7960-7970.

9- Appendix

Appendix I

Chapter 2: Sample PCR reaction mixtures and programs

Sample PCR reaction mixture

PCR kit	Phusion® High-Fidelity DNA Polymerase (M0530L, New England Biolabs)	Expand High fidelity (HIFI, 11732 641001, Roche)/Taq polymerase
	Volume (µl)	Volume (µl)
Water	33.5	38.5
Buffer	10 (5x buffer)	5 (10x buffer)
DNA	1	1
Primers (F+R)	2.5 of 100 µM each	2.5 of 100 µM each
dNTPs	1	1
DMSO	1.5	1.5
Enzyme	0.5	0.5
Total volume	50	50

Standard Phusion PCR program

Step	Temperature (^{o}C)	Time (seconds)	cycles
Initial denaturation	98	30	1
Denaturation Annealing Extension	98 56 72	12 30 60/kb	25-35
Final extension	72	700	1
Hold	4	-	-

Standard HIFI or Taq DNA polymerase PCR program

Step	Temperature (^{o}C)	Time (seconds)	cycles
Initial denaturation	94	120	1
Denaturation Annealing Extension	94 56 72	15 30 60/kb	25-35
Final extension	72	720	1
Hold	4	-	-

PCR program used to amplify selection cassette from AdZ using HIFI

Step	Temperature (^{o}C)	Time (seconds)	cycles
Initial denaturation	94	120	1
Denaturation Annealing Extension	94 56 72	30 30 270	9
Denaturation Annealing Extension	94 56 72	30 30 270 (+20 per successive cycle)	24
Final extension	72	420	1
Hold	4	-	-

Appendix II

Chapter 2: Primer and oligonucleotide sequences

First round primers used to amplify Merlin genes

	F	R
RL8A	AAGACACCGGGACCGATCCAGCCTGGATCCTTTTCCTTTCTTCTCCGTTTCTCC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGCTAAAAACAGCGGACAGTCC
RL9A	AAGACACCGGGACCGATCCAGCCTGGATCCCGCGCCGCCGGTCCCCC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGAGAAACAGCACGTAGGTCAGG
UL7/8	ACACCGGGACCGATCCAGCCTGGATCCACGTGGTAAATAAATCATGG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCAGTTCCGTGTCTGTCATAAACA
UL36	GACACCGGGACCGATCCAGCCTGGATCCGCGAAGGTAGAGGAGTCCGTCATGGA	GTTGTTCATGTAGGCGTGTGGCGGAGCGGGTTAGGGATTGGCTTACCAGCGCT
UL50	AAGACACCGGGACCGATCCAGCCTGGATCCTCGGCGGTGGCGTCGGTGCG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGTCGCGGTGTGCGGAGCGTG
UL69	AAGACACCGGGACCGATCCAGCCTGGATCCGCGACGGCGGCCATGGAGCT	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGTCATCCATATCATCGCTGTAACAC
UL73	AAGACACCGGGACCGATCCAGCCTGGATCCTCGTGGCGGTGGTGTGATGG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTATAGCCTTTGGTGGTGGTCG
UL77	AAGACACCGGGACCGATCCAGCCTGGATCCGTTCTGGACTATCTGGGACG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCAACACCGCCACGCTCGGA
UL82	AAGACACCGGGACCGATCCAGCCTGGATCCCTGACCCTCCCCCCATCCC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGATGCGGGGTCGACTGCGTG
UL83	AAGACACCGGGACCGATCCAGCCTGGATCCGCCGCTCAGTCGCCTACAC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTACCTCGGTGCTTTTTGGGC
UL84	AAGACACCGGGACCGATCCAGCCTGGATCCGCCCTCTCGCGCCCGCAG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGAGATCGCCGCAGACCATG
UL85	AAGACACCGGGACCGATCCAGCCTGGATCCTCTCCGCTTCTCCTCCGTC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGCCTTTAAATATGCAGGTCGCGG
UL93	AAGACACCGGGACCGATCCAGCCTGGATCCTTCCGAATATGGAAACGCACC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTAAGATCGTCGAACGGCAAGC
UL95	AAGACACCGGGACCGATCCAGCCTGGATCCGAATCTCCGCGCAACATGATG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTTAGATTCAACGTGATGAGACCCG
UL128	AAGACACCGGGACCGATCCAGCCTGGATCCCGCGCGTCATGAGTCCCA	GAGCGGGTTAGGGATTGGCTTACCAGCGCTATAGGGAACGCTGCCAGCGG
UL130	AAGACACCGGGACCGATCCAGCCTGGATCCGCCTGCGTCACGGGAAATAA	GAGCGGGTTAGGGATTGGCTTACCAGCGCTAACGATGAGATTGGGATGGGTG
UL132	AAGACACCGGGACCGATCCAGCCTGGATCCATCCACGTCGCCACGTCTCG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGTCGTACTCGGGATCTCTGAGCG
UL148	AAGACACCGGGACCGATCCAGCCTGGATCCTTCCTAACCCGCGCAGCATG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCCGACGCCGCGACACCAGGT
UL147A	AAGACACCGGGACCGATCCAGCCTGGATCCGACTGCGCTGGTGATCATGTC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGATCACACAAGTGACGAGGAGCG
UL147	AAGACACCGGGACCGATCCAGCCTGGATCCGCAGAAAATATCATGTTTTCAATATGTTGC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCCAGCGCAGTCTAAAGTAGTGGTAA
UL146	AAGACACCGGGACCGATCCAGCCTGGATCCCCGGATATTACGAATTATTGGTAGTGAC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTTTCCTTCAGACCTACTAGGGTTACAG
UL145	AAGACACCGGGACCGATCCAGCCTGGATCCGTCACTCTTTCTTGTTGCCTTCG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTATCTTCACTTCCACCCATCGGG
UL144	AAGACACCGGGACCGATCCAGCCTGGATCCTCTTCCGGTAGGAGGCATGAAG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCAGGGTGCGGTAGAAAATTTTGC

UL142	AAGACACCGGGACCGATCCAGCCTGGATCCGCAGTGTGTGAAATGTGAATAGTGTG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCTGACCGCGCCATACTTCGT
UL140	AAGACACCGGGACCGATCCAGCCTGGATCCCTCTTCAGACTTGTCATGACCCC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCAGGGTCTGATGAAGCTGCCAA
UL139	AAGACACCGGGACCGATCCAGCCTGGATCCCGACAAAGGGGTTCGTATATCAAT	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCCGAGGCGGAGGTGGAAATG
UL138	AAGACACCGGGACCGATCCAGCCTGGATCCGACAGAGGACGGTCACCATGG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCGTGTATTCTTGATGATAATGTACCATGG
UL148A	AAGACACCGGGACCGATCCAGCCTGGATCCCGTCGGCGGCGATAGCCATG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGTAACACTCGTCCGACACTTCCAC
UL148B	AAGACACCGGGACCGATCCAGCCTGGATCCGGGAGCGAACGGGATGGC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTAAGTCGCGTGTGGTAGATCTCCT
UL148C	AAGACACCGGGACCGATCCAGCCTGGATCCCCTTCTCCGCGAGCAAATGTT	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGGGGAAGGCTGTAAGCACGG
UL148D	AAGACACCGGGACCGATCCAGCCTGGATCCGAGGAGTCGCGGCATGACG	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGCCGCTGCCGCTCCC
US23	AAGACACCGGGACCGATCCAGCCTGGATCCCCTCTTTTGACATGTGGCGTACA	GAGCGGGTTAGGGATTGGCTTACCAGCGCTCACAAAGTGCTCCCGAAAATCGA
US24	AAGACACCGGGACCGATCCAGCCTGGATCCGCTGAGAGTGTCGCGCAATGATGGAT	GAGCGGGTTAGGGATTGGCTTACCAGCGCTAATCTGGATGTACTCGCGCACACCCG
US33A	AAGACACCGGGACCGATCCAGCCTGGATCCGGCGCAAAAGCACCATGAGC	GAGCGGGTTAGGGATTGGCTTACCAGCGCTGGACCGCGGCACGTAAAACG

Second round primers for amplifying Merlin genes

F	AACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTGGATCC
R	GGCGTGACACGTTTATTGAGTAGGATTACAGAGTATAACATAGAGTATAATATAGAGTATACAATAGTGACGTGGGATCC

Adz sequencing primers

From MIE Promoter	AATGTCGTAACAACTCCG
From MIE PolyA	ACCTGATGGTGATAAGAAG

Ad recombinant internal primers: where sequencing from MIE promoter and MIE polyA was insufficient to cover the full length of the insert, internal primers were designed)

US23F1	CCACAGACGCCGTTTGCTAC
US23F2	CTACGAACACGGTCTGCGGC
UL36F1	CCGTCTGTTCGCAAGGTAAGC
UL36F2	CGCCTTCTGGACCCACAACAC
UL36F3	TGACGAGGATGACGACGACG
UL69F1	CTTCTCACCCCTCCTCAGCC
UL69F2	TATCTGGAGACGGTGGGCGG
US24F1	CACTCTGCGGAGAATGACGG

Sequencing primers for modified Merlin BACs:

UL4-tags-seq-F	GTATGGAATAACGAATCTCACG
UL4-tags-seq-R	GAAACGCAAACCGACCAAC
UL74A-V5seq-F	TCCGTCTGCCTTGTGGTTTG
UL74A-V5seq-R	TCTACGCGCTATCGGCCATC
UL150A-V5seq-F	CGTCGGCTTCCAGCTTGTCT
UL150A-V5seq-R	CTCAGCGGTCGCCCTATCAC

Amplifying selection cassette for insertion into Merlin BAC: (insert target sequence):

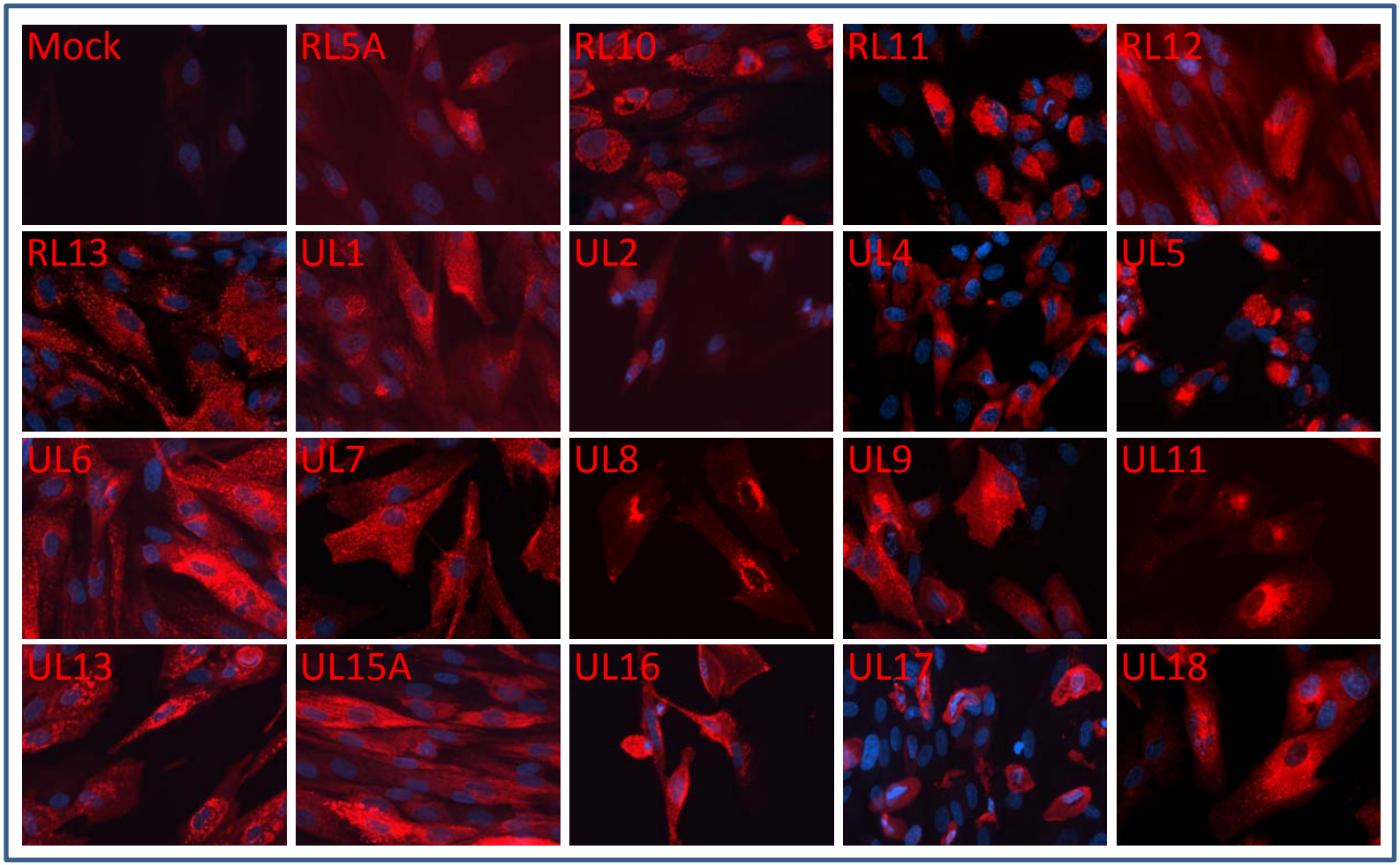
UL4-L	AACAGCGGAAAGTATTATTTCAAACGAGAAGATGTGAATTCCACCTTTTATTACTCTTGTTACAACCTGACCGTGACCCCTGTGACGGAAGATCACTTCG
UL4-R	CGTAGGGACGCTTAACATTTCAAAAGCAACGTAAACACAATAGCTACAGCCACGCGGCTCTGTGGAACTTTACATGCGTCTGAGGTTCTTATGGCTCTTG
UL74A-L	CTAGTCTCCCGCACCTTGGCTGTGTACCGAGCTTGTATACAGGGCCCGATATGTAACCAAACCCACAACAGTACCTCGCCTGTGACGGAAGATCACTTCG
UL74A-R	GGGCACTCTTGCTCATTGTAATAAAGTACACGCCACACGGTGTGATGATACTATATGCGTGAGGTTTGTGCGTCTTTATCTGAGGTTCTTATGGCTCTTG
UL150A-L	TTCTCTCGTGGAGGCGACGGCGAATGCAGCAGACGGTGTTCGATGGAGATGGCGTGCGAGGAAGAAAGCGCCGTGTTGCCTGTGACGGAAGATCACTTCG
UL150A-R	GACAAGCACAAGTGGACACCGCCATCTGCCACCACACATGGCCCATGTGCTCCGACGTCCCGCATCCAACGTCGTCTGCCTGAGGTTCTTATGGCTCTTG

Oligonucleotides to replace selection cassette from Merlin BAC with C-terminal V5 tag:

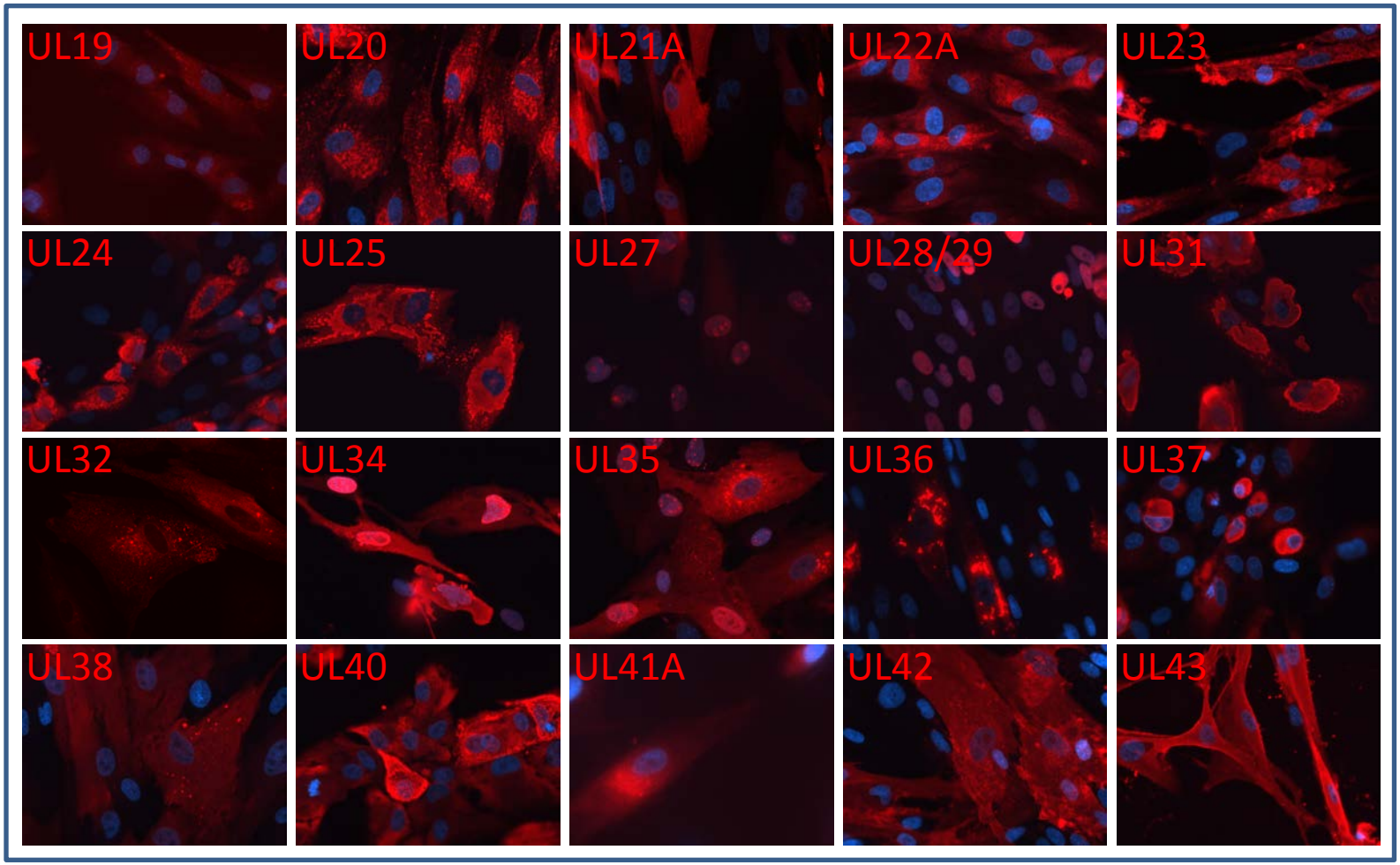
UL4-V5-F	AGATGTGAATTCCACCTTTTATTACTCTTGTTACAACCTGACCGTGACCAGCGCTGGTAAGCCAATCCCTAACCCGCTCCTAGGTCTTGATTCTACGTAA
UL4-V5-R	GTAAACACAATAGCTACAGCCACGCGGCTCTGTGGAACTTTACATGCGTTTACGTAGAATCAAGACCTAGGAGCGGGTTAGGGATTGGCTTACCAGCGCT
UL74A-V5-F	AGCTTGTATACAGGGCCCGATATGTAACCAAACCCACAACAGTACCTCGAGCGCTGGTAAGCCAATCCCTAACCCGCTCCTAGGTCTTGATTCTACGTAA
UL74A-V5-R	CGCCACACGGTGTGATGATACTATATGCGTGAGGTTTGTGCGTCTTTATTTACGTAGAATCAAGACCTAGGAGCGGGTTAGGGATTGGCTTACCAGCGCT
UL150A-V5-F	CAGACGGTGTTCGATGGAGATGGCGTGCGAGGAAGAAAGCGCCGTGTTGAGCGCTGGTAAGCCAATCCCTAACCCGCTCCTAGGTCTTGATTCTACGTAA
UL150A-V5-R	ACCACACATGGCCCATGTGCTCCGACGTCCCGCATCCAACGTCGTCTGCTTACGTAGAATCAAGACCTAGGAGCGGGTTAGGGATTGGCTTACCAGCGCT

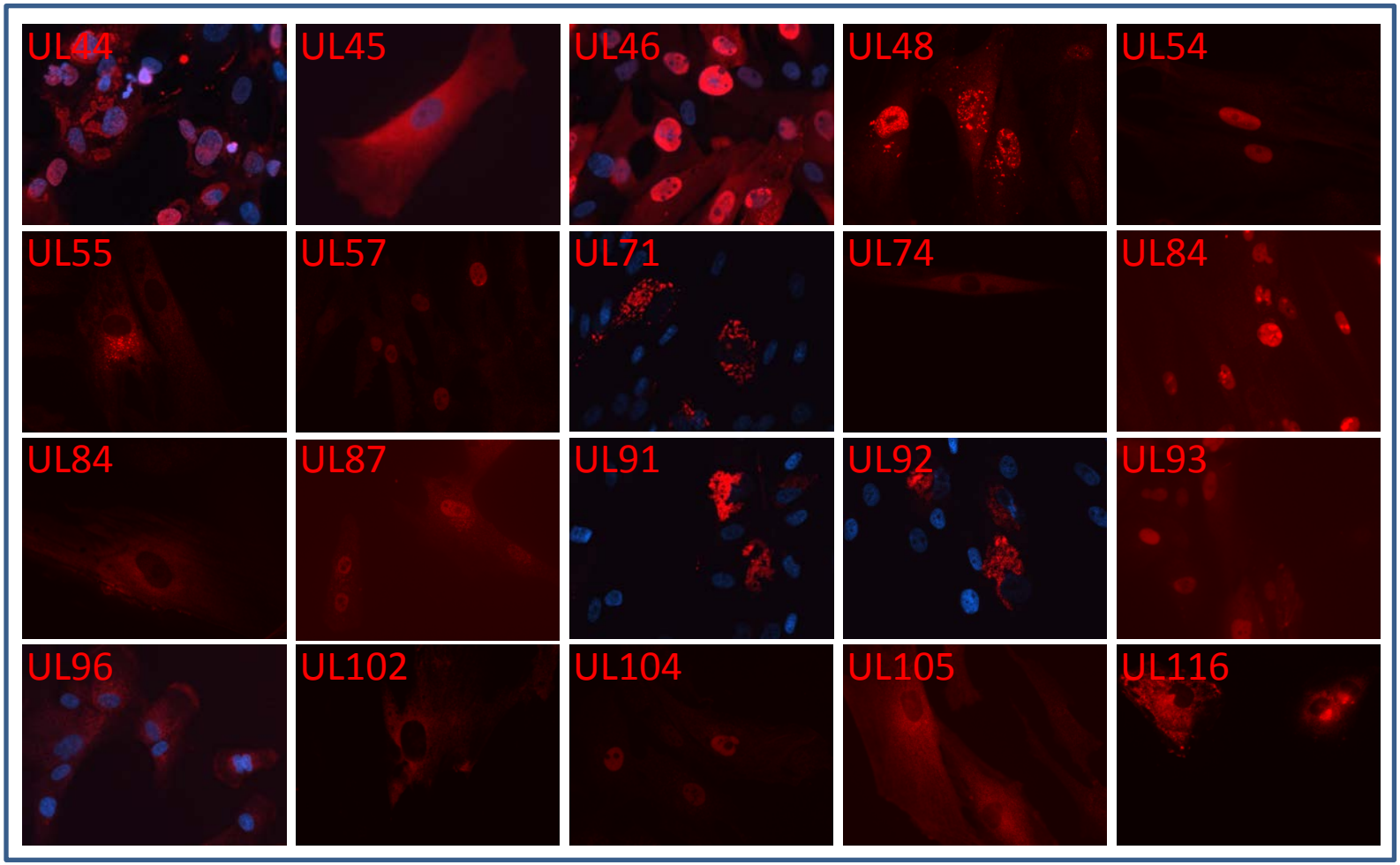
Appendix III

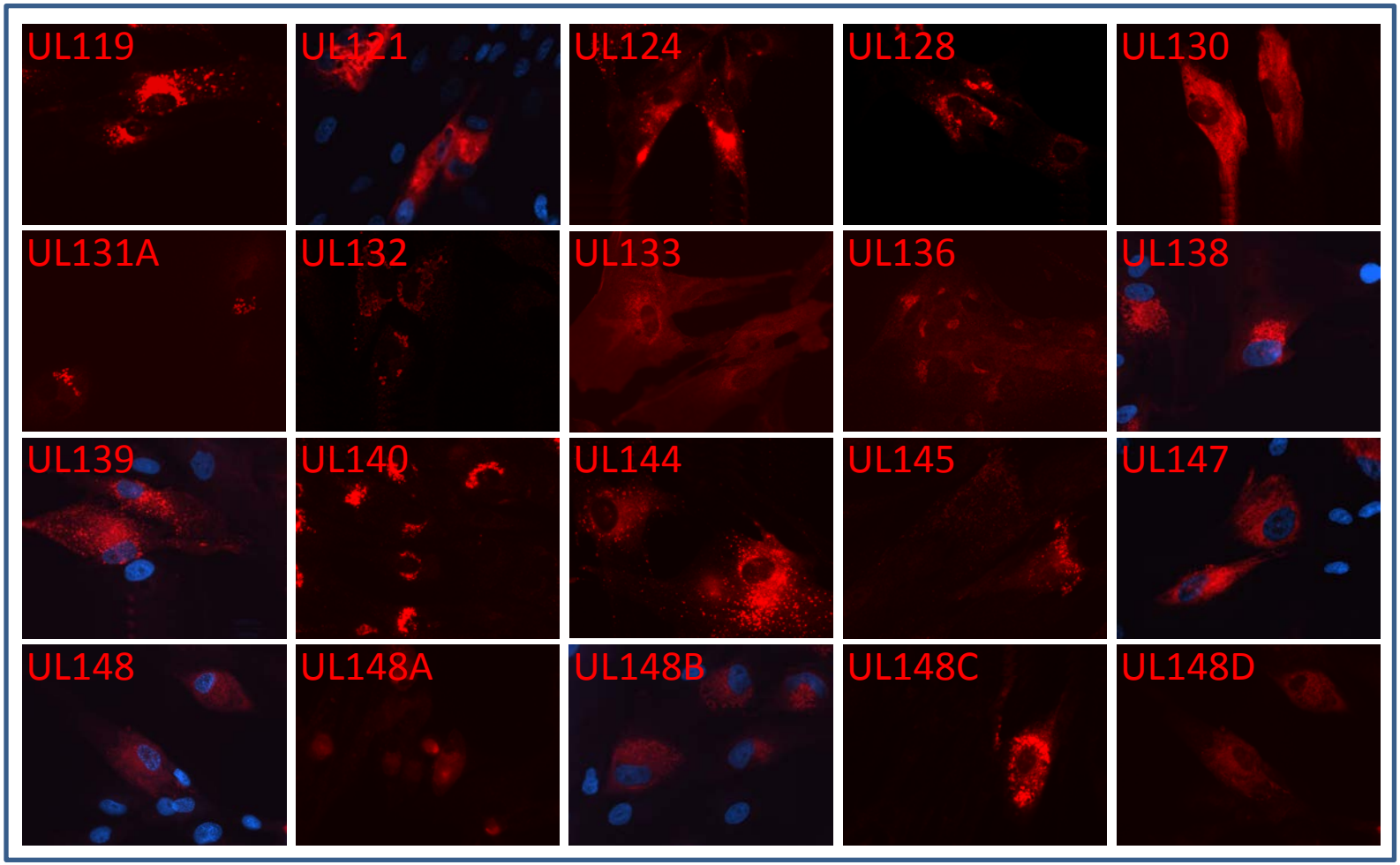
Chapter 3: Immunofluorescence images lacking DAPI and/or Phalloidin co-staining

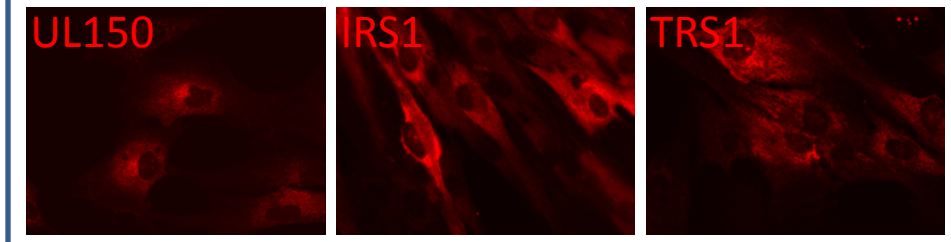

Mock
RL5A
RL10
RL11
RL12
RL13
UL1
UL2
UL4
UL5
UL6
UL7
UL8
UL9
UL11
UL13
UL15A
UL16
UL17
UL18

x


UL19
UL20
UL21A
UL22A
UL23
UL24
UL25
UL27
UL28/29
UL31
UL32
UL34
UL35
UL36
UL37
UL38
UL40
UL41A
UL42
UL43


UL44
UL45
UL46
UL48
UL54
UL55
UL57
UL71
UL74
UL84
UL84
UL87
UL91
UL92
UL93
UL96
UL102
UL104
UL105
UL116


UL119
UL121
UL124
UL128
UL130
UL131A
UL132
UL133
UL136
UL138
UL139
UL140
UL144
UL145
UL147
UL148
UL148A
UL148B
UL148C
UL148D


UL150
IRS1
TRS1

Appendix IV

Publications

Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication

Richard J. Stanton,[1] Katarina Baluchova,[2] Derrick J. Dargan,[2] Charles Cunningham,[2] Orla Sheehy,[2] Sepehr Seirafian,[1] Brian P. McSharry,[1] M. Lynne Neale,[1] James A. Davies,[1] Peter Tomasec,[1] Andrew J. Davison,[2] and Gavin W.G. Wilkinson[1]

[1]Section of Medical Microbiology, Tenovus Building, School of Medicine, Cardiff University, Cardiff, United Kingdom. [2]MRC Virology Unit, Institute of Virology, University of Glasgow, Church Street, Glasgow, United Kingdom.

Human cytomegalovirus (HCMV) in clinical material cannot replicate efficiently in vitro until it has adapted by mutation. Consequently, wild-type HCMV differ fundamentally from the passaged strains used for research. To generate a genetically intact source of HCMV, we cloned strain Merlin into a self-excising BAC. The Merlin BAC clone had mutations in the RL13 gene and UL128 locus that were acquired during limited replication in vitro prior to cloning. The complete wild-type HCMV gene complement was reconstructed by reference to the original clinical sample. Characterization of viruses generated from repaired BACs revealed that RL13 efficiently repressed HCMV replication in multiple cell types; moreover, RL13 mutants rapidly and reproducibly emerged in transfectants. Virus also acquired mutations in genes UL128, UL130, or UL131A, which inhibited virus growth specifically in fibroblast cells in wild-type form. We further report that RL13 encodes a highly glycosylated virion envelope protein and thus has the potential to modulate tropism. To overcome rapid emergence of mutations in genetically intact HCMV, we developed a system in which RL13 and UL131A were conditionally repressed during virus propagation. This technological advance now permits studies to be undertaken with a clonal, characterized HCMV strain containing the complete wild-type gene complement and promises to enhance the clinical relevance of fundamental research on HCMV.

Introduction

Human cytomegalovirus (HCMV) is a clinically important herpesvirus that is ubiquitous in human populations worldwide (1). Primary infection is followed by lifelong persistence, during which virus reactivation must be constrained continuously by host immune surveillance. Myeloid cell progenitors are a recognized site of latency, with infectious virus being produced following differentiation into macrophages or dendritic cells. Severe disease is most commonly observed when immunity is compromised by infection (e.g., HIV/AIDS) or immunosuppressive therapy (e.g., in transplant recipients). HCMV is also the leading viral cause of congenital disability and malformation, which was the primary basis for it being designated a highest priority level vaccine target by the US Institute of Medicine (2). HCMV disease can manifest as a wide range of clinical conditions (e.g., pneumonia, colitis, retinitis, hepatitis, arteriosclerosis, or systemic infection), reflecting the capacity of the virus to infect a wide range of cell types in vivo (3–7). Despite this wide tropism in vivo, only fibroblasts support the efficient growth of cultured strains in vitro.

Appreciation of HCMV pathogenesis is heavily dependent on research performed using the high-passage strains AD169 and Towne. However, these strains are known to have both lost virulence and experienced substantial alterations in their genomes during passage in cell culture (8–10). The HCMV genome may be represented as *ab*-U_L-*b′a′c′*-U_S-*ca*, where U_L and U_S denote long and short unique regions and *ba*/*b′a′* and *ca*/*c′a′* indicate inverted repeats flanking the unique regions. AD169 and Towne have acquired large deletions (15 kb and 13 kb, respectively) in the U_L/*b′* region in addition to defects elsewhere in genes that are dispensable for growth in fibroblasts (11–18). The U_L/*b′* region encodes a viral CXCL chemokine (gene UL146; ref. 19), a tumor necrosis factor receptor homolog (gene UL144; ref. 20), natural killer cell evasion functions (genes UL141 and UL142; refs. 21–23), a regulator of latency (gene UL138; refs. 24, 25), and several other uncharacterized functions. Therefore, the U_L/*b′* region is likely to contribute to virulence via several pathways. Spontaneous defects clearly arise elsewhere in the HCMV genome, yet their significance is only occasionally defined functionally (13, 18, 26, 27). Indeed, whole-genome sequencing of multiple clinical strains following growth in vitro reveals that clinical HCMV genomes invariably mutate whether passaged in fibroblasts, epithelial cells, or endothelial cells (28).

There is a need to develop and characterize low-passage HCMV strains for experimental applications. However, the generation of laboratory strains of WT HCMV is problematic not only because clinical samples often contain multiple strains (27, 29–36), but especially because genetic adaptations occur even during the early stages of passage in cell culture. Thus, most strains passaged in fibroblasts are mutated in 1 of 3 adjacent genes (UL128, UL130, and UL131A; collectively termed the UL128 locus, UL128L) whose products form a complex bound to glycoproteins gH and gL in the virion envelope (37–43). These mutations inhibit formation of the complex and thereby render the virus incapable of infecting epithelial, endothelial, and certain myeloid cell types (12, 44–48).

Authorship note: Richard J. Stanton and Katarina Baluchova contributed equally to this work.

Conflict of interest: The authors have declared that no conflict of interest exists.

Citation for this article: *J Clin Invest.* 2010;120(9):3191–3208. doi:10.1172/JCI42955.

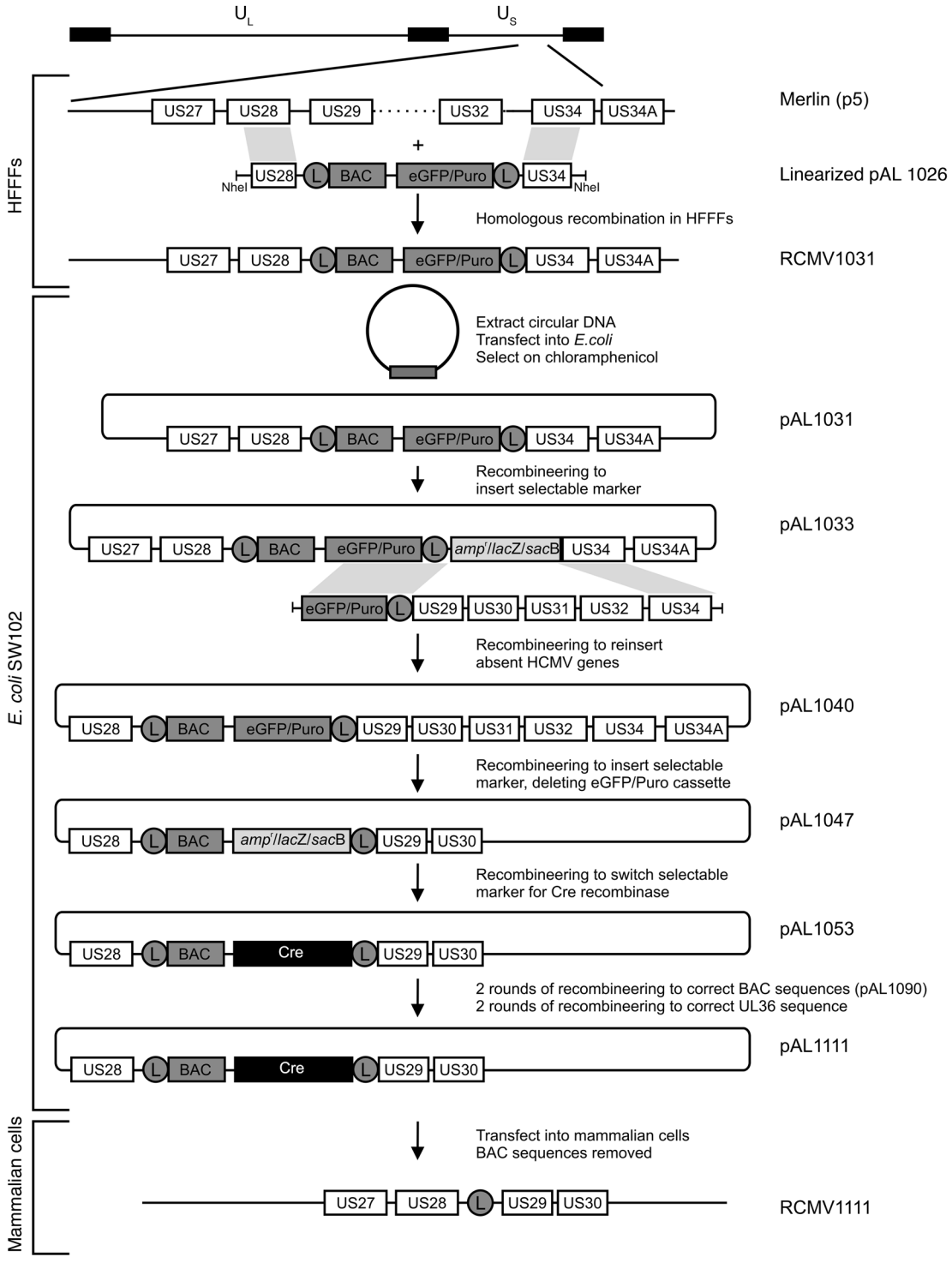


Figure 1
Steps in the construction of pAL1111. The approximate location of the insertion site of the BAC vector between US28 and US29 in the U_S region of the HCMV genome is shown at the top. The designations of BACs (pAL series) and viruses (RCMV series) are shown on the right. Boxes indicate protein-coding regions (labeled), and circles denote *loxP* sites (L). The BAC box represents pBeloBAC11, the eGFP/Puro box represents a cassette expressing eGFP and a puromycin resistance protein, and the Cre box represents a Cre recombinase gene containing a synthetic intron.

Table 1
Sequence differences between Merlin (NC_006273) and the HCMV genome in pAL1053

Position	Change (NC_006273 → pAL1053)	Coding region affected	Coding effect
11363	CAAAAAAAC → CAAAAAAAAC	RL13	*f*
40231	TTGGAGG → TTGCAGG	UL32	None
48828	GGCGAAC → GGCAAAC	UL36	R → C
49054	CATGATC → CATCATC	UL36	I → M
49203	GACGGTA → GACAGTA	UL36	R → C
49791	AGGGGTG → AGGAGTG	UL36	*P* → S
49792	GGGGTGC → GGGCTGC	UL36	H → Q
194592–194643 (1025–1076)	ACACCCCCCGTCCCCGCACACCCCCCGTCCCCGCACACCCCCCGTCCCCG → G	None (*b*/*b*′)	None
195063 (605)	GGACCGC → GGAGCGC	None (*b*/*b*′)	None
194781	CCGTAGA → CCGCAGA	None (*b*′ only)	None

Nt positions are given with respect to NC_006273. *f*, frameshift. Altered bases underlined.

The genomes of several HCMV strains have been captured as BACs, including a number based on low-passage viruses, and these are proving to be an invaluable resource for research (49–54). However, none contains the full complement of HCMV genes, both because of mutations that occurred in cell culture prior to being cloned and, in most cases, because sequences were deleted in order to accommodate the BAC vector. In addition, the original clinical material is not available for most of these strains, thus preventing the sequences of the cloned genomes being verified against those of unpassaged virus. To provide a reliable source of HCMV gene sequences and a reproducible source of genetically intact, clonal virus for pathogenesis studies, we sought to produce an infectious BAC containing the complete HCMV gene complement. HCMV strain Merlin (ATCC VR-1590) was selected because it is designated the reference HCMV genome sequence by the National Center for Biotechnology Information (GenBank AY446894; RefSeq NC_006273) and the original clinical sample is available for comparison. Merlin was derived from the urine of a congenitally infected infant and sequenced after 3 passages (p3) in human fibroblasts (27).

Our study commenced with the construction of a BAC clone containing the complete genome of strain Merlin (p5). DNA sequencing led to the identification of disabling mutations in genes RL13 and UL128, which were repaired first singly and then in combination to produce BACs containing a complete WT gene complement. Experiments conducted with viruses derived from the BACs revealed that RL13 expression severely impairs HCMV replication in fibroblast and epithelial cell cultures. Thus, when viruses generated from BACs containing the WT RL13 gene were passaged in vitro, further mutants emerged rapidly, producing a phenotypically altered virus with enhanced growth properties. Although RL13 is known to be a member of a restricted set of HCMV hypervariable genes (27), its expression had not, to our knowledge, previously been characterized. Use of the BACs showed that RL13 encodes a highly glycosylated protein located in the virion envelope. Since virus generated from the WT BAC was susceptible to mutation, cell lines were constructed in which expression of RL13 and UL131A could be modulated. For what we believe is the first time, these cell lines permit experiments to be undertaken using a characterized HCMV strain containing the complete WT gene complement.

Results

Cloning the HCMV strain Merlin genome. We aimed to clone the complete HCMV strain Merlin genome by inserting a BAC vector into a noncoding region between virus genes US28 and US29. Preliminary studies (data not shown) indicated that straightforward insertion of the BAC-targeting vector into the HCMV genome by homologous recombination during virus growth in primary human fetal foreskin fibroblasts (HFFFs) was associated with compensating deletions in virus sequences, presumably due to genome size constraints operating during virus DNA packaging. The BAC-targeting vector was therefore redesigned to replace the region of the HCMV genome containing US29–US34 (Figure 1). BAC vector DNA was transfected into HFFFs, which were superinfected with HCMV strain Merlin (p5), and puromycin selection was used to enrich for recombinants. Circular DNA was extracted and electroporated into *E. coli*. A total of 22 clones were analyzed by restriction endonuclease digestion (data not shown), and all contained the HCMV genome minus US29–US34. Since both U_L and U_S may be present in either orientation in an HCMV genome, a linear molecule will be one of 4 isomers and a circular molecule one of 2. Restriction endonuclease profiles corresponding to both circular conformations were detected among the BAC clones (data not shown). A single clone (pAL1031; Figure 1) of the 12 clones with U_L and U_S present in the standard arrangement was selected for further analysis.

Recombineering (55, 56) was used to insert the US29–US34 region into pAL1031, thus generating pAL1040, which contains the entire Merlin genome (Figure 1). Cre/lox technology has been deployed previously to promote excision of prokaryotic vector sequences (also located between US28 and US29) from an AD169 BAC following transfection into HFFFs (53). We adopted this strategy to produce a self-excising Merlin BAC by 2 rounds of recombineering, in which the enhanced GFP (eGFP)/Puro cassette in pAL1040 was replaced by a gene encoding Cre recombinase, thus generating pAL1053 (Figure 1). The version of Cre used contained a synthetic intron to prevent its expression in *E. coli* (57, 58). Following transfection into HFFFs however, Cre recombinase was expressed and mediated removal of the BAC vector by recombination between *loxP* sites engineered at the junctions with the virus genome. Thus, the only exogenous sequences remaining in virus generated from pAL1053 and subsequent BACs were those of single *loxP* and NheI sites (40 bp in total) located between US28 and US29 (Figure 1).

Sequence of BAC pAL1053. Sequencing of the entire HCMV component of pAL1053 confirmed that the whole Merlin genome had been captured and that a single *a*/*a*′ sequence was present between

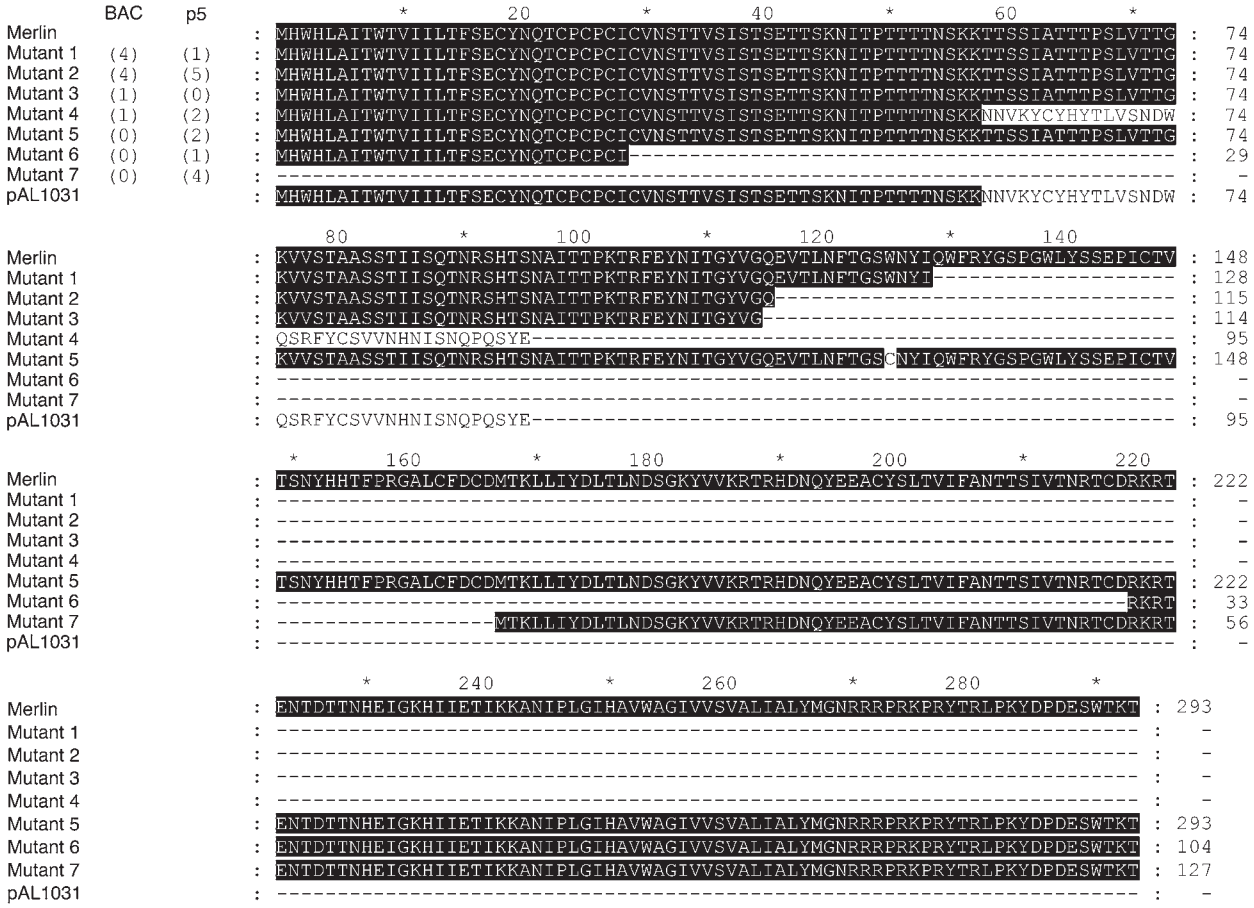


Figure 2

Alignment of amino acid sequences potentially encoded by mutated versions of RL13 in pAL1031 and 10 other BACs from this stage, and PCR clones from Merlin (p5). Shading indicates identity to the WT Merlin sequence (NC_006273). Numbers in parentheses indicate the number of clones containing the relevant mutation. Nt mutations were as follows (numbering relative to strain Merlin sequence, NC_006273): mutant 1, C → T (nt 11573; premature stop codon); mutant 2, G → T (nt 11534; premature stop codon); mutant 3, C → T (nt 11531; premature stop codon); mutant 4 (which includes pAL1031), extra A (nt 11363; frameshift); mutant 5, G → C (nt 11563, W → C); mutant 6, deletion of nt 11276–11860; mutant 7, G → A (nt 11191, M → I, altered start codon; the protein is depicted as starting from next in-frame M).

the *b* and *c*, and *b′* and *c′*, sequences. As in Merlin (p3), pAL1053 contained a G to A (G → A) substitution at nt 176311 that introduces an in-frame stop codon and would lead to premature termination of the UL128 protein. pAL1053 differed from Merlin (p3) at a total of 12 locations (Table 1). These included 5 substitutions clustered in UL36. The sequences of PCR products amplified from the original clinical material, Merlin (p3), the earliest BAC precursor of pAL1053 (pAL1031), and 10 additional Merlin BACs from the same stage as pAL1031 showed that these mutations were present only in pAL1031, whereas the sequences from all other sources were identical to that of the Merlin (p3) sequence. Thus, the differences in UL36 were atypical and limited to the Merlin genome that was captured in pAL1031. The UL36 mutations in pAL1053 were repaired by recombineering to match the Merlin (p3) sequence, thus generating pAL1111 (Figure 1).

The most significant additional difference between Merlin (p3) and pAL1053 (and hence pAL1111) was a frameshift in RL13 at nt 11363 caused by a single nt insertion (Table 1; also see below). The remaining differences had no apparent effect on protein-coding potential. A synonymous substitution in gene UL32 was found retrospectively to represent a single nt polymorphism in the Merlin (p3) genome. Three alterations were noted in the *b*/*b′* inverted repeat sequence. Two of these were substitutions, one being present in both *b* and *b′* and the other in *b′* alone. The other was a 51-bp deletion due to natural length variation in a tandem repeat sequence in both *b* and *b′*. The *b*/*b′* sequence is the region of the HCMV genome that is most prone to variation in HCMV (18, 59)

Mutations in RL13. To investigate the origin of the frameshift mutation in the RL13 coding region (nt 11189–12070), this locus was sequenced in 10 additional BACs from the same stage as pAL1031. Surprisingly, all were found to contain disruptive mutations, and 4 different classes (classes 1–4) of mutant were identified, each of which is predicted to express a truncated RL13 protein (Figure 2). To determine whether RL13 was mutated in the virus stock used to generate pAL1031, the coding region was amplified by PCR from Merlin (p5) and 15 clones were sequenced (Figure 2). Mutations that were detected in single clones, except where they corresponded to mutations in the BACs, could have arisen by PCR error and were excluded from the analysis. The PCR clones identified mutations that corresponded to 3 of the 4 mutant groups

Table 2
Sequence differences between RL13 in WT Merlin (NC_006273) and subpopulations of Merlin (p3) represented in bacteriophage M13 or PCR clones

Position	Mutation (NC_006273 → clone)	Residue	Coding effect	No. of M13 clones	No. of PCR clones
11191	CATGCAC → CATACAC	1	M → I	1/6	0/22
11537	GAAGTGA → GAAATGA	117	V → M	4/8	11/22
11563	ATGGAAT → ATGCAAT	125	W → C	2/8	1/22
11563	ATGGAAT → ATGCAAT	125	W → C	2/8	4/22
11900	ATTGAAA → ATTCAAA	238	E → Q	2/8	4/22
11884	AATCGGG → AATTGGG	232	None	1/8	2/22
11964	TATCGGT → TATAGGT	259	S → X	1/8	2/22

Nt positions are given with respect to NC_006273. X, termination codon. Altered bases underlined.

(groups 1, 2, and 4) in BACs, plus 3 additional groups (classes 5–7). None of the BAC or PCR clones yielded a sequence representing the WT RL13 coding region.

The finding that all genomes in Merlin (p5) and BACs derived from Merlin at the pAL1031 stage were mutated in RL13 raised the question of at which stage these mutations had arisen. The original Merlin sequence was determined from bacteriophage M13 clones derived from Merlin (p3). Retrospective analysis of the sequence database revealed that, although the RL13 consensus was WT, all individual M13 clones were mutated at various different positions (Table 2). Moreover, an analysis of individual PCR clones produced from Merlin (p3) was also consistent with this virus stock comprising a mixture of RL13 mutants (Table 2). Thus, 18/22 PCR clones exhibited substitutions that were also detected in the M13 clones. Each of the remaining 4/22 PCR clones contained unique mutations in RL13 (data not shown), though it is not possible to exclude the possibility that these resulted from PCR errors.

To investigate whether the mutants detected in Merlin (p3 and p5) were present in the clinical sample or whether they originated during cell culture, 10 PCR clones of RL13 were generated from the original clinical sample, using the same primers and conditions as were used for Merlin (p5). Nine clones were WT in sequence, and one clone contained a single synonymous substitution (A → G at nt 11959). This substitution was not detected in any versions of Merlin from passage in cell culture and thus most probably resulted from a PCR error. These data indicated that RL13 was predominantly WT in the clinical sample and are consistent with RL13 mutants having arisen and been selected during cell culture.

Growth properties of virus generated from the Merlin BAC. Virus (RCMV1111) was readily recovered from HFFFs transfected with pAL1111 (Figure 1) and exhibited a restriction endonuclease profile identical to that of Merlin (p5) (Figure 3A). A pattern diagnostic of the presence of all 4 genome isoforms was distinguishable in the linear RCMV1111 genome. Infections of HFFFs at low MOI (0.01 PFU/cell) demonstrated that RCMV1111 grew with kinetics comparable to those of Merlin (p5) (Figure 3B).

In generating RCMV1111, the BAC vector was excised from the genome to leave only a residual sequence of 40 bp between US28 and US29. Intracellular FACS indicated that RCMV1111 expressed the US28 protein during the course of productive infection at an abundance comparable with Merlin (p5) (Figure 3C). Since an antibody was not available for the US29 protein, US29 transcript levels expressed by RCMV1111 and Merlin (p5) were measured by quantitative RT-PCR (QRT-PCR) (Figure 3D), using UL123 (IE1) as a standard, and were comparable. Thus, the residual 40 bp in RCMV1111 had no discernible effect on expression of the adjacent US28 and US29 genes.

Expression of eGFP by virus generated from the Merlin BAC. To facilitate the characterization of viruses generated from Merlin BACs, an IRES-eGFP expression cassette was inserted into pAL1111 immediately downstream from UL122 (IE2), so that expression was under the control of the major IE promoter (60), generating pAL1158. In the resulting virus (RCMV1158), eGFP was expressed with IE kinetics (data not shown). In growth studies performed at low MOI (0.01 PFU/cell) in HFFFs, RCMV1158 exhibited a small but consistent reduction in levels of virus production (2-fold) at 12 days post-infection (PI) (Figure 3E).

Repair of RL13 or UL128 results in growth defects in fibroblasts. In order to reconstitute the complete HCMV gene complement, the lesions in UL128 and RL13 in pAL1111 were repaired seamlessly by recombineering, both singly and in combination, culminating in a set of 4 BACs (pAL1111, pAL1119, pAL1120, and pAL1128; Table 3). The complete sequence of the fully repaired BAC (pAL1128) was verified by Solexa sequencing. A corresponding set of eGFP-tagged BACs was also constructed (pAL1158, pAL1159, pAL1160, and pAL1161; Table 3).

To assess the efficiency of virus spread from cell to cell, the 4 eGFP-tagged BACs were transfected individually into HFFFs, the cells were overlaid, and then plaque areas were measured at 3 weeks post-transfection (PT) (Figure 4 and Figure 5A). Comparable results were obtained using the BACs lacking the eGFP tag (data not shown). Viruses with either RL13 (RCMV1159) or UL128 (RCMV1160) repaired produced much smaller plaques than the double mutant (RCMV1158). Furthermore, the virus (RCMV1161) with both genes repaired produced even smaller plaques. RL13 and UL128 thus act independently to restrict either cell-to-cell transmission or virus production in fibroblast cultures.

To investigate further, HFFFs transfected with the eGFP-tagged BACs without overlay were monitored over time. Plaques formed initially by viruses with UL128, RL13, or both genes repaired (RCMV1159, RCMV1160, and RCMV1161) were subsequently outgrown by uninfected cells. To encourage virus dissemination, cells transfected with pAL1159, pAL1160, or pAL1161 were trypsinized at 7-day intervals and reseeded in the same flask without the addition or removal of cells. The eGFP reporter gene was used to identify infected cells (Figure 5B), and cell-free virus levels were measured by titration (Figure 5C). The rates of virus spread through the monolayer (Figure 5B) were consistent with the

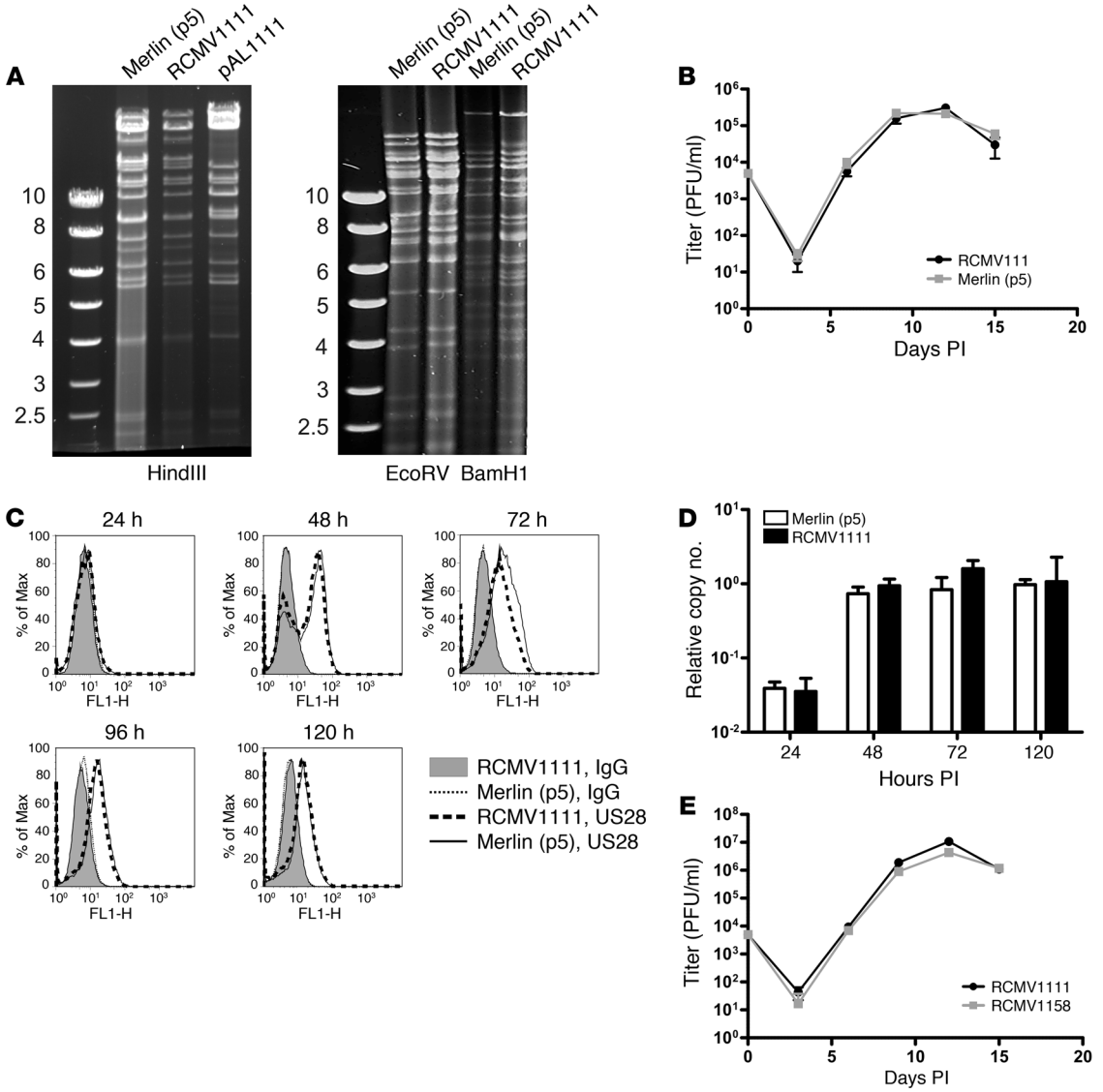


Figure 3

Properties of Merlin BAC–derived virus RCMV1111. (**A**) Restriction endonuclease profiles of DNA extracted from Merlin (p5), RCMV1111 (the virus derived from pAL1111), and pAL1111. Markers (kb) are shown. (**B**) Virus titers released into the medium following infection of HFFFs at an MOI of 0.01 PFU/cell. Error bars show mean ± SD; results are representative of 2 experiments. (**C**) Intracellular FACS staining for the US28 protein in cells at various times PI with either Merlin (p5) or RCMV1111. Controls in the form of samples stained with IgG are shown. (**D**) Copy numbers of US29 transcripts relative to UL123 (IE1) transcripts at various times PI with Merlin (p5) or RCMV1111. Figures were averaged from 3 experiments; error bars show mean ± SD. (**E**) Virus titers released into the medium following infection of HFFFs at an MOI of 0.01 PFU/ml with RCMV1158 or RCMV1111. Error bars show mean ± SD; results are representative of 2 experiments.

relative plaque sizes (Figure 5A), with slowest growth occurring when RL13 and UL128 were both repaired (Figure 5, A–C). Repairing either RL13 or UL128 also gave a clear reduction in the levels of infectious virus released into the supernatant (Figure 5C). Repairing both genes delayed dissemination of the infection through the monolayer, but substantial and increasing amounts of infectious virus were released by 5–7 weeks PT (Figure 5C).

In titrations conducted to generate Figure 5C, the plaque sizes were largely in accord with those reported in Figure 5A. However, some plaques from the repaired viruses were substantially larger than the majority. Plaque sizes were measured at the earliest time point possible: 2 weeks PT for RCMV1158, 3 weeks PT for RCMV1159 and RCMV1160, and 4 weeks PT for RCMV1161 (Figure 5D). Two clearly distinct plaque sizes were evident for RCMV1159. The 3 larger plaques were expanded individually in HFFFs, and RL13 was amplified by PCR and sequenced. Each virus contained the same 2 mutations in RL13, specifically C → A at nt 11286 (resulting in S → Y) and C → G at nt 11436 (resulting in the introduction of an in-frame stop codon). Sequencing of viruses derived from the 3 largest RCMV1161 plaques did not reveal any mutations in RL13 or UL128L.

Repairing RL13 results in growth defects in epithelial cells. Restoration of RL13 independently of UL128 suppressed virus replication in fibroblasts, with approximately 10% of virus having remutated in

Table 3
WT (+), mutated (–) status of RL13 and UL128 in Merlin BACs

BAC	Derived virus	RL13	UL128
BACs containing no tags			
pAL1111	RCMV1111	–	–
pAL1119	RCMV1119	+	–
pAL1120	RCMV1120	–	+
pAL1128	RCMV1128	+	+
BACs tagged with eGFP			
pAL1158	RCMV1158	–	–
pAL1159	RCMV1159	+	–
pAL1160	RCMV1160	–	+
pAL1161	RCMV1161	+	+
BACs containing RL13 tagged with a V5 epitope			
pAL1279	RCMV1279	+	–
pAL1280	RCMV1280	+	+
BACs tagged with eGFP, tet Operators before RL13/UL128L			
pAL1393	RCMV1393	–	+ (tet0)
pAL1448	RCMV1448	+ (tet0)	–
pAL1498	RCMV1498	+ (tet0)	+ (tet0)

RL13 by 3 weeks PT (see above; Figure 5D). To determine whether this effect was limited to fibroblasts or whether RL13 may be advantageous in other cell types (as is the case for UL128L), retinal pigmented epithelial cells (ARPE-19s) were transfected with pAL1160 or pAL1161 (Table 3) and the size of plaques was measured (Figure 6A). At all time points measured, plaques from the virus lacking RL13 (pAL1160) were larger than those from the virus in which RL13 had been repaired (pAL1161). Therefore, RL13 conferred a growth restriction in epithelial cells as well as fibroblasts. In contrast, UL128L is known to confer a restriction only in fibroblasts.

Cell cultures transfected with the eGFP-tagged BACs were monitored simultaneously for the number of infected cells (Figure 6B) and amount of infectious virus released into the supernatant (Figure 6C). The lack of growth of UL128 mutants (RCMV1158 and RCMV1159) in ARPE-19s (data not shown) was consistent with previous observations that UL128L is essential for epithelial cell tropism. As was the case with HFFFs, plaques generated by viruses in which UL128 was repaired (RCMV1160 and RCMV1161) were eventually overgrown by uninfected cells, and virus dissemination required periodic reseeding of the monolayer. Both RCMV1160 and RCMV1161 grew much more slowly in ARPE-19s than HFFFs, and, unlike HFFFs, ARPE-19s did not attain complete infection (Figure 6B), instead reaching a plateau and forming a chronic infection. The virus in which both UL128 and RL13 were repaired (RCMV1161) grew much more slowly than RCMV1160 and formed a chronic infection in which a much lower proportion of cells was infected (approximately 15% for RCMV1161 compared with 35% for RCMV1160). RCMV1161 also produced lower levels of cell-free virus than RCMV1160 (Figure 6C). This effect was only partially attributable to the lower numbers of infected cells, since RCMV1161 infected about half the number of cells as did RCMV1160 and yet produced over 10-fold less cell-free virus. Comparable results were obtained using viruses lacking the eGFP marker. Thus, repairing RL13 results in slower cell-to-cell spread and reduced levels of cell-free virus from epithelial cells as well as fibroblasts.

Rapid generation and selection of RL13 and UL128L mutants in fibroblasts. To generate working stocks of BAC-derived viruses in HFFFs, the cells in a 25-cm^2 flask were transfected with BAC DNA and reseeded periodically until 100% were infected. Cell-free virus was transferred to a single 150 cm^2 flask and then to five 150 cm^2 flasks, and virus stocks (p3) were prepared. To test the genetic integrity of the virus, 10 PCR clones each of RL13 and UL128L were sequenced. Differences from the original sequences were excluded as probable PCR errors if they were detected in single clones. An analysis of 13 virus stocks (Table 4), each derived independently from BAC DNA, found that (a) when only RL13 was repaired, it mutated in all stocks (stocks 1–4); (b) when only UL128 was repaired, the UL128 locus mutated in 2 of 3 HFFF derived stocks (stocks 5, 6), whereas it remained intact in epithelial derived stocks (stocks 8, 9; a synonymous substitution was detected in one instance); and (c) when both RL13 and UL128 were repaired, RL13 mutated in 1 stock (stock 11), whereas the UL128 locus appeared to remain intact. When the stock (stock 11) containing an RL13$^+$UL128$^+$ virus that harbored a frameshift in RL13 in 20% of genomes was passaged a fourth time, the RL13 mutation was then observed in 100% of clones, along with a deletion compromising both UL128 and UL130 in the UL128 locus. Thus, both RL13 and the UL128 locus remutated in fibroblasts, with the former tending to mutate more rapidly.

RL13 encodes a virion envelope protein. Since RL13 has not previously been characterized, it was important to establish and investigate its expression. RL13 was therefore tagged with a sequence encoding a C-terminal V5 epitope and inserted into a recombinant adenovirus (RAd) vector. The V5 epitope was also fused to RL13 within the strain Merlin BACs, in both a UL128$^-$ (RCMV1279) and a UL128$^+$ (RCMV1280) background (Table 3). The primary translation product of RL13 is predicted to be a 35-kDa protein containing a signal sequence, a transmembrane domain, 7 potential N-linked glycosylation sites, and 26 potential O-linked glycosylation sites. When RL13V5 was expressed in isolation using the RAd vector, 80-kDa and 55-kDa proteins were detected (Figure 7A), whereas 100-kDa and 55-kDa species were detected in cells infected with RCMV1279 (RL13V5$^+$UL128$^-$; Figure 7B). The 55-kDa proteins were susceptible to EndoH digestion, indicative of them being an ER-retained immature form, whereas the 80- and 100-kDa proteins were resis-

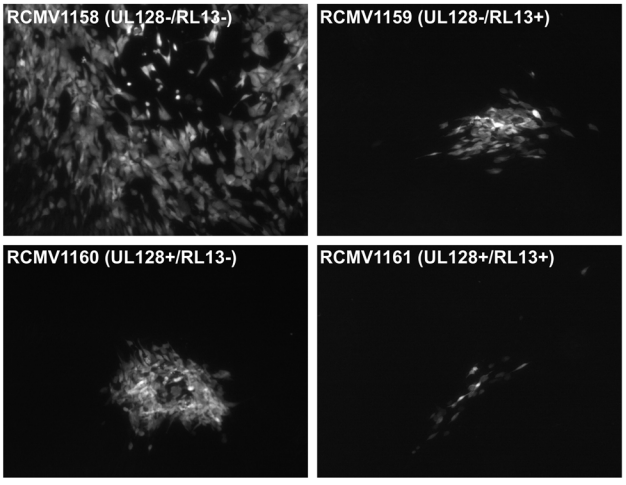


Figure 4
Growth of BAC-derived viruses in HFFFs. Images show eGFP expressed by plaques formed in HFFFs 3 weeks PT with BAC DNA as indicated. Scale bars: 100 μm.

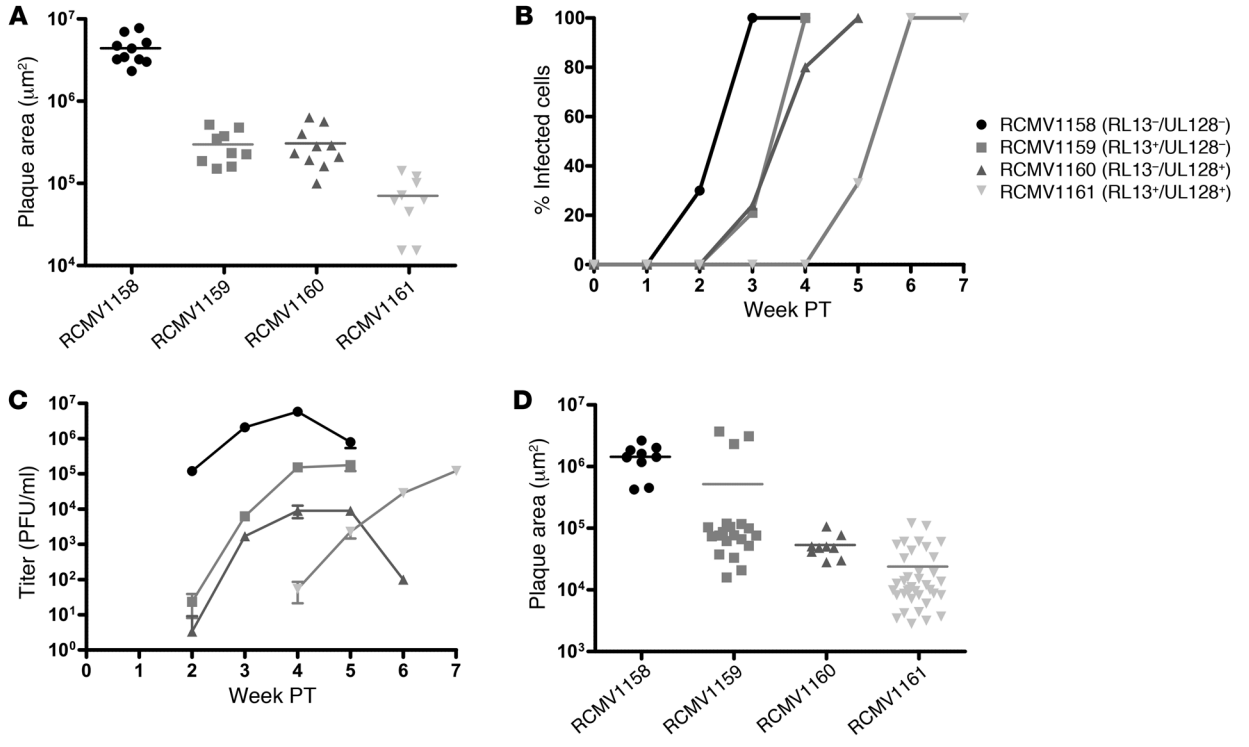


Figure 5
Growth of BAC-derived viruses in HFFFs. (**A**) Areas of individual plaques at 3 weeks PT, with cells overlaid to prevent cell-free spread of virus. (**B**) Kinetics of infection measured by FACS analysis. (**C**) Kinetics of infectious virus release into the medium of the cultures in **B**, with error bars showing mean ± SD. (**D**) Area of plaques generated by virus harvested from the media in **C**, with each sample derived from the earliest time point at which more than 10 plaques were visible.

tant to EndoH digestion and are thus presumably fully mature. PNGaseF digestion reduced the 55-kDa proteins to 35 kDa and, in addition, reduced the 80- and 100-kDa proteins to 57 and 65 kDa, respectively. Since neither enzyme was able to reduce the 80- and 100-kDa proteins to 35 kDa, it is likely that mature gpRL13 contains O-linked, as well as N-linked, sugars. Identical results were obtained with cells infected with RCMV1280 (RL13V5$^+$UL128$^+$), and no bands were seen using lysates from RL13$^-$ virus (data not shown). The observation that gpRL13 appears to be more extensively glycosylated in the context of HCMV infection than when expressed in isolation is consistent with previous observations on CD155 and gpUL18, to the effect that HCMV infection alters normal cellular glycosylation processes (22, 61).

The V5 epitope was also used to track gpRL13 expression by immunofluorescence. When RL13V5 was expressed using the RAd vector, gpRL13 trafficked to discrete cytoplasmic vesicles, a proportion of which could be stained with the early endosomal marker Rab5A (Figure 7C). This observation is consistent with much of the protein having transited the Golgi apparatus. In the context of a productive HCMV infection, the intracellular distribution of gpRL13 was different. The protein localized to the cytoplasmic site of virion assembly and colocalized in both HFFFs and ARPE-19s with a marker for the trans-Golgi network (TGN46) and pp28 (an outer tegument protein encoded by gene UL99 that interacts with the envelope and is acquired by the virion in the cytoplasm; Figure 7D) and gH (a virion envelope glycoprotein encoded by gene UL75; Figure 7E).

The finding that gpRL13 is a glycoprotein that tracks to the site of virion assembly raised the possibility that it might be a virion surface envelope protein. Growth of RL13$^+$ HCMV in vitro is inefficient, and spontaneous mutants will arise and be rapidly selected (see above). Nevertheless, we found that limited, short-term growth of RL13$^+$ virus could be fostered by sequential reseeding of cultures (Figure 5), and the antibody to a C-terminal tag would preferentially recognize full-length, nonmutated gpRL13 even if a proportion of mutants had arisen during passage. In virions purified from such cultures, only the 100-kDa mature, Endo H–resistant form of gpRL13 was detected. In virion fractionation studies, this protein copurified with glycoprotein B (gB; the product of gene UL55) in the soluble envelope fraction, rather than with pp65 (the product of gene UL83) in the tegument (Figure 8A). The V5 epitope is predicted to be located topologically on the inner side of the envelope, and this orientation was supported by immunoelectron microscopy, in which gold-labeled secondary antibody exhibited extensive labeling around the inner surface of the envelope (Figure 8B).

Conditional expression of RL13 and UL128L. The BAC pAL1111 (RL13$^-$UL128$^-$) has utility as a reproducible source of fully characterized, clonal HCMV that exhibits good genetic stability and rapid growth to high titers in fibroblasts. We also show above that short-term experiments can be performed using the Merlin BAC in which RL13 and UL128 are repaired. The major phenotypic impact these genes have on HCMV biology emphasized the need to work with WT virus. However, the rapid emergence of mutants at these loci during amplification from Merlin BAC

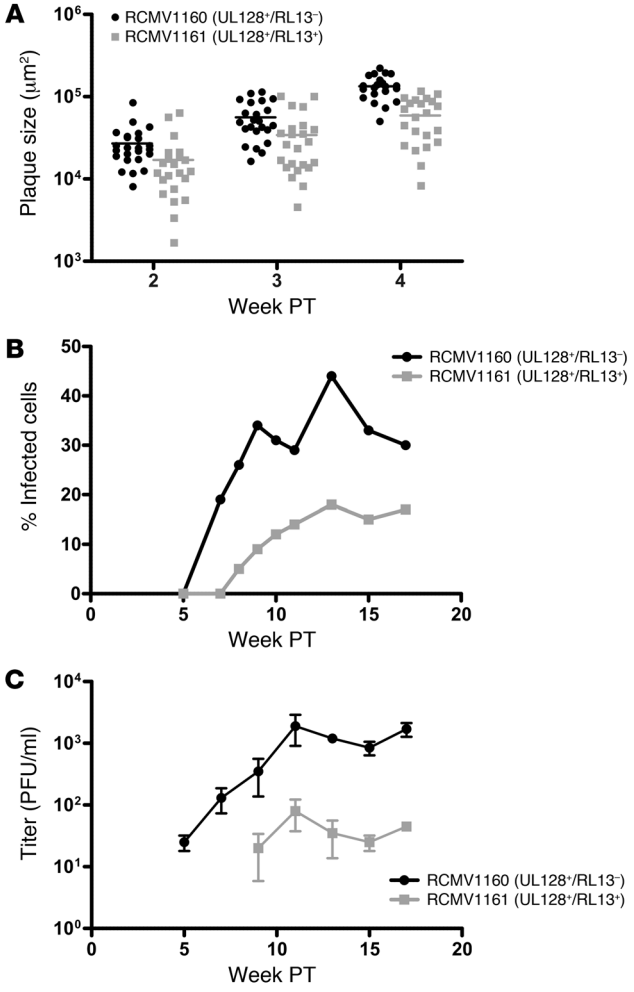


Figure 6
Growth of BAC-derived viruses in ARPE-19s. (**A**) Size of plaques at 2, 3, or 4 weeks PT, with cells overlaid to prevent cell-free spread of virus. (**B**) Kinetics of infection following transfection of BACs into ARPE-19s without overlay, measured by FACS analysis of eGFP. (**C**) Kinetics of infectious virus release into the medium of the cultures in **B**, with error bars showing mean ± SD.

transfections meant virus stocks were inevitably contaminated with mutants. In order to enable studies with phenotypically WT virus, we sought to suppress expression of RL13 and the UL128 locus during virus expansion following Merlin BAC transfection. To achieve this, hTERT-immortalized HFFFs were transduced with a retrovirus encoding the tet repressor containing a C-terminal nuclear localization signal (NLS) and tet operators were inserted upstream of the translation initiation codon of the gene to be modulated (62, 63). In HFFF-tet cells, the tet repressor binds to the tet operators and prevents transcription of the target gene, whereas transcription proceeds as normal in parental HFFFs. When tested with a RAd-expressing tet-regulated GFP, expression levels in HFFF-tet cells were reduced 180-fold relative to parental HFFFs, showing little more fluorescence than cells infected with an empty vector control (Figure 9A).

The transcriptional initiation sites for RL13 are located approximately 30- and 110-bp upstream from the translational initiation codon (K. Baluchova, unpublished observations). Transcription of UL128 starts within the UL130 coding region, and an mRNA initiating approximately 10 bp upstream of UL131A potentially encodes both the UL131A and UL130 proteins (26, 43). tet operators were inserted in various locations upstream from the RL13 or UL131A coding regions, and their capacity to selectively promote replication of RL13$^+$ or UL128$^+$ viruses in HFFF-tet cells was tested empirically. The optimal location for the tet operators was determined to be a single tet operator 19 bp upstream of RL13 (pAL1448) and 2 tet operators 33 bp upstream of UL131A (pAL1393; Table 3). Plaque formation was inhibited approximately 10-fold in HFFF cells by the presence of an intact UL128 locus, and this inhibition was relieved by infection of HFFF-tet cells with a tet-controlled UL131A virus, RCMV1393 (Figure 9, B and C). Similarly, plaque formation was inhibited approximately 10-fold in HFFF cells by the presence of an intact RL13, and this inhibition was relieved by infection of HFFF-tet cells with a tet-controlled RL13 virus, RCMV1448 (Figure 9, D and E). Titers of tet-controlled viruses obtained from HFFF-tet cells were increased 18- to 32-fold relative to the non-tet–controlled viruses (Figure 9F). In addition, sequencing of 10 PCR clones of UL128L and RL13 from RCMV1393 and RCMV1448, respectively, showed all contained the WT sequence. Following passage in HFFFs with this RCMV1393 stock, virus was able to infect ARPE19 cells and produced plaques of a size equivalent to the non–tet-controlled parental virus RCMV1160 (Figure 9G). Finally, having verified that both RL13 and UL131A could be independently tet-controlled, a virus was constructed in which both genes contained tet operators upstream of their ATG codons. Like the previous constructs, plaque formation was inhibited in HFFF cells by the presence of an intact RL13 and UL128 locus, and this repression was relieved by infection of HFFF-tet cells with a virus in which both RL13 and UL128 were tet-controlled (RCMV1498) (Figure 10, A and B).

Thus, repression of RL13 and UL131A in HFFF-tet cells demonstrably provides a means to amplify Merlin virions containing intact versions of RL13 and the UL128 locus. Two options are available to produce phenotypically WT HCMV from these viruses: either transcriptional repression can be released by the addition of doxycycline to infected HFFF-tet cells or, as demonstrated above, a final replication cycle can be performed in HFFFs.

Discussion

Diagnostic laboratories have long recognized that HCMV strains in clinical samples do not adapt readily to cell culture. Our construction of what we believe is the first BAC containing a complete, characterized copy of a genetically intact HCMV genome provides an explanation for this phenomenon in showing that adaptation is dependent on 2 independent mutations, one in the UL128 locus (previously recognized; ref. 26) and one in RL13 (identified in the present study). To date, HCMV research has by necessity used viruses that, to varying degrees, are compromised genetically. Although laboratory-adapted viruses have proved invaluable, there is a clear need to develop systems that represent the WT virus responsible for clinical disease. To this end, the Merlin genome (p5) was cloned using BAC technology, and the sequence of an initial BAC (pAL1053) was compared with that of the parental strain. Specific issues relating to tissue culture adaptation were clarified and resolved using sequence data derived directly from the original clin-

Table 4
RL13 and UL128L mutations detected by PCR in BAC-derived viruses at p3

Stock	RL13 mutations				UL128L mutations			
	Mutation	Position	Affected clones (%)	Effect	Mutation	Position	Affected clones (%)	Effect
RL13$^+$UL128$^-$ BAC transfected into HFFFs								
1	^A	11363	33	*f*				
1	G → T	11901	40	E → X				
2	ΔA	11363	30	*f*				
3	^A	11363	50	*f*				
4	A → G	11297	20	S → G				
4	C → T	11427	20	A → V				
4	A → G	11956	20	I → M				
RL13$^-$UL128$^+$ BAC transfected into HFFFs								
5					C → T	176229	14	R → Q
5					T → C	176845	14	I → V
5					CAAGA → TCTTG	176915-9	21	FL → SA
6					ΔT	176456	30	None (intron)
7					None		100	None
RL13$^-$UL128$^+$ BAC transfected into ARPE-19s								
8					None		100	None
9					A → G	176917	20	None
RL13$^+$UL128$^+$ transfected into HFFFs								
10	None		100		ΔT	176456	50	None (intron)
11	^A	11363	20	*f*	None		100	None
12	None		100		ΔT	176456	30	None (intron)
13	None		100		ΔT	176456	30	None (intron)

Each row represents an individual virus stock. Nt positions are given with respect to NC_006273. +, WT gene; –, mutated gene; ^, inserted nt; Δ, deleted nt.

ical sample that gave rise to Merlin. Mutations that affected HCMV coding regions were sequentially repaired by recombineering, so that the only relic in virus recovered from BACs by DNA transfection was a 40-bp sequence containing a *loxP* and an NheI site located between the US28 and US29 coding regions. The presence of this sequence did not affect expression of the flanking genes.

The high level of sequence identity between pAL1053 and strain Merlin (p3) demonstrated that the HCMV genome had suffered minimal perturbation during BAC cloning. In pAL1053, nonsynonymous mutations in protein-coding regions were identified in UL36, UL128, and RL13. The differences in UL36 were specific to pAL1053 and its predecessor, pAL1031, and were presumably derived from a small proportion of mutants that arose in cell culture. HCMV and rhesus cytomegalovirus strains have previously been found to have acquired inactivating mutations in UL36 following passage in vitro, implying that there may be some selective pressure for such mutants (14, 64). Following repair of UL36, the resulting BAC (pAL1111) provided a reproducible source of clonal HCMV strain Merlin (RCMV1111; RL13$^-$UL128$^-$; Table 3), whose growth properties were indistinguishable from the parental virus. BACs in which either RL13, UL128, or both genes were repaired were generated from pAL1111 and supplemented by a corresponding series of BACs containing an eGFP marker. pAL1128 and pAL1161 contain the complete HCMV gene complement and, for what we believe is the first time, enabled experiments to be undertaken with a clonal population of effectively WT virus. Although RCMV1111 and its eGFP-expressing counterpart (RCMV1158) were stable during limited culture and suitable for generation of high-titer virus stocks, repair of UL128 or RL13 (or both) severely impaired HCMV growth in HFFFs, and propagation of these viruses was associated with rapid selection of further mutants at these loci. Therefore, we developed the HFFF-tet cell line, which allowed RL13 or UL131A to be repressed during recovery of virus from the BAC, thus providing a stable platform for experimentation on genetically intact HCMV.

Although HCMV replication has been described in a wide range of cell types, efficient virus production in vitro has been achieved only using viruses adapted to growth in fibroblasts. Indeed, existing cell-culture systems appear to be essentially incompatible with efficient propagation of WT HCMV. HCMV is not exceptional in this regard. Of the 8 known human herpesviruses, only WT herpes simplex virus types 1 and 2 (HSV-1 and HSV-2) can be propagated readily to high titer in cell culture. Nonetheless, very high HCMV virus loads are often detected in urine samples from congenitally infected neonates (as was the case for the sample from which Merlin was isolated), and this indicates that conditions exist in vivo that are compatible with efficient replication of WT HCMV. It is possible that RL13 regulates a switch to productive HCMV replication in vivo.

The UL128 locus is essential for the efficient infection of epithelial and endothelial cells by HCMV (37–43) but is detrimental to growth in fibroblasts. Consequently, mutants in this locus are selected when clinical isolates are passaged in fibroblasts (12, 26, 27, 44, 45, 47, 48). Merlin was originally selected for detailed investigation from a series of clinical isolates partly on the basis of its excellent growth properties in fibroblasts. However, this process favored the selection of a UL128 mutant, which emerged during p1 (26). When the lesion was repaired in the BACs, alternative defects in the UL128 locus emerged by p3–p4. In other studies monitoring endothelial cell tropism in clinical isolates, loss of phenotype has been reported in fibroblasts after p20 (44), p22 (45), or p15–p20 (28). Such studies defined a virus "passage" as corresponding to

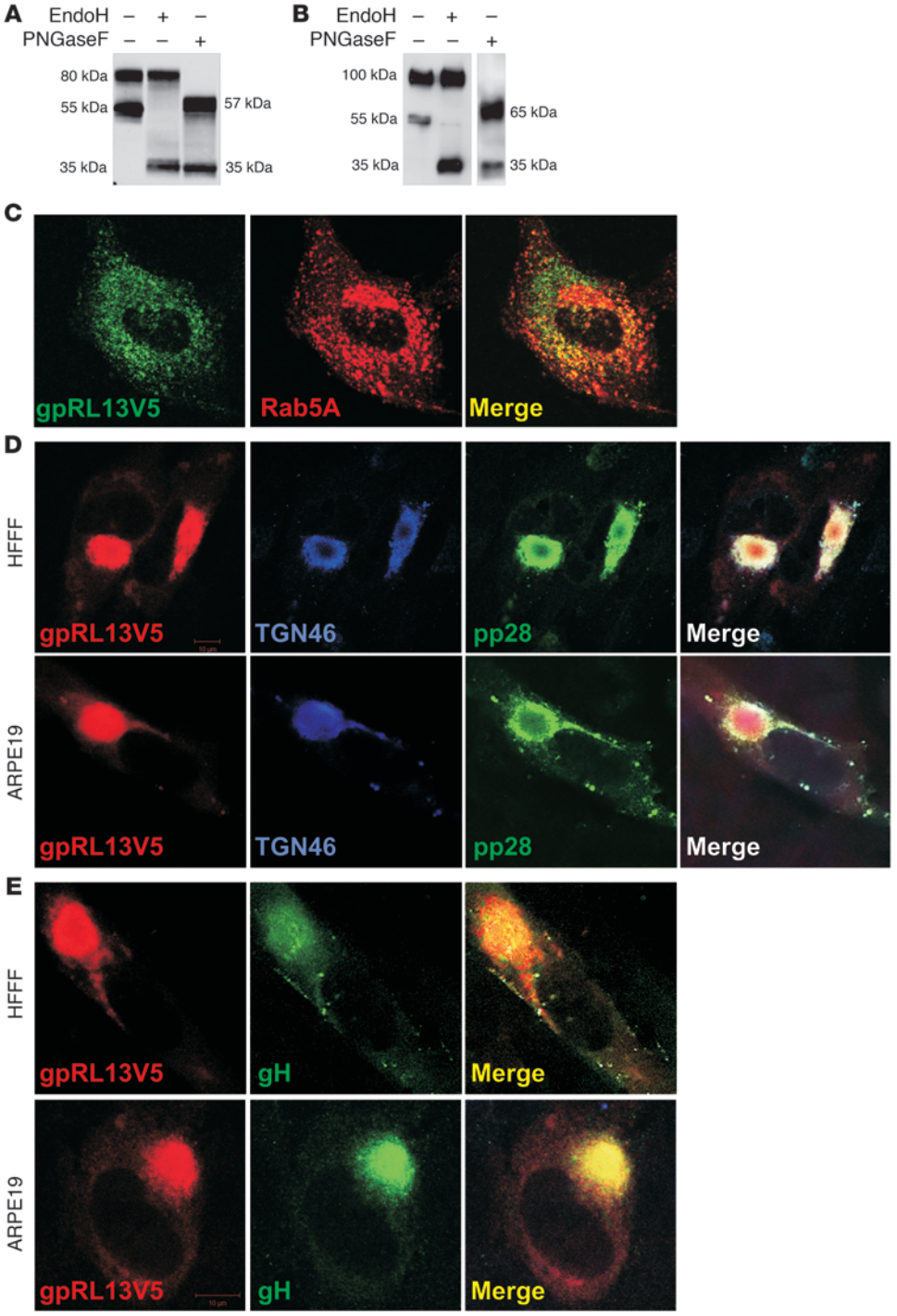


Figure 7
Characterization of gpRL13. (**A** and **B**) Western blot performed on RL13V5, showing the sizes of the proteins in native form or following digestion with EndoH or PNGaseF. RL13V5 was either expressed in isolation from a RAd (**A**) or from its native position in RCMV1279 (**B**). (**C**) Immunofluorescence performed for the indicated antigens on HFFFs infected with the RAd expressing RL13V5 (×640 magnification). (**D** and **E**) Immunofluorescence performed for the indicated antigens either on HFFFs infected with RCMV1279 or on ARPE-19s infected with RCMV1280. Original magnification, ×640.

each serial subculture (usually weekly) of an infected cell monolayer. However, in the present study, virus passage involved destruction of the entire cell monolayer following infection with cell-free virus, a process that normally involves many more cycles of virus replication. Also, given that viruses mutated in the UL128 locus released higher amounts of virus into the medium, our use of cell-free virus may have further encouraged the selection of mutants.

The identification of an RL13 mutation in pAL1053 was unexpected because the genome sequence of Merlin (p3) indicated that this coding region was intact (27). It led to the discovery that low-passage Merlin stocks actually consisted of mixtures of various different RL13 mutants. This finding brings a resolution with observations that almost all HCMV strains and BACs are overtly mutated in RL13 (27). Parallels exist between RL13 and the UL128 locus in the fact that mutations arise in both genes during adaptation to culture in fibroblasts. However, it is clear that RL13 and the UL128 locus can function independently because either gene in intact form was capable of suppressing virus replication, and

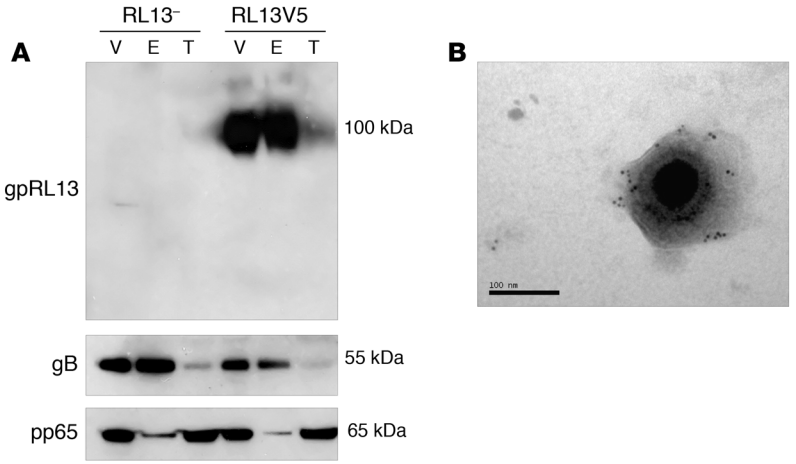


Figure 8
Localization of RL13 in the virion. (**A**) Western blot performed on purified virions from virus expressing a truncated gpRL13 (RCMV1111, RL13–) or full-length RL13 with a C-terminal epitope tag, gpRL13V5 (RCMV1279, RL13+). Complete virions (V), envelope fraction (E), and tegument/capsid fraction (T) were probed for the antigens indicated. (**B**) Immunogold EM staining for RL13V5 in purified RCMV1279 virions. Original magnification, ×40,000.

repair of both had an additive effect. The RL13 phenotype was also shown to be distinct from that of the UL128 locus, since RL13 suppressed HCMV replication in both fibroblasts and epithelial cells, whereas the UL128 locus was not only stable, but promoted infection, in epithelial cells. Nevertheless, the finding that gpRL13 is present in the virion indicates that it may have a role in modifying tropism (similar to the UL128 locus) or in modulating cell signaling during virus entry (as proposed for gB).

The apparent ease with which RL13 mutants were selected in cell culture raised the possibility that they preexisted in the clinical sample, perhaps reflecting the potential for an expanded cell tropism in vivo. This possibility is technically challenging to disprove, since such mutants might be present in very low proportions. However, we were unable to detect the existence of RL13 mutants in the primary clinical sample from which Merlin was derived. It was estimated that approximately 10^6 particles were used in the initial infection of a 25-cm^2 flask to isolate Merlin, and this would have resulted in most cells receiving a particle. Nonetheless, passage 1 of Merlin took 4 weeks to complete. The mutation rate of DNA viruses is relatively low and has been calculated at 0.003 per genome replication for HSV-1 (65). If this value holds for HCMV, mutants in the UL128 locus and/or RL13 would be expected to arise during p1. This mutation rate operating on 10^6 viruses would result in a total of 3,000 mutations per genome replication. Since RL13 accounts for 879/235,646 bp in strain Merlin genome, approximately 11 mutations would be in RL13. The acquisition of mutations in both RL13 and UL128 could have resulted either from sequential mutations in the same template or as a result of recombination between independent mutants. Despite the counterintuitive nature of such rapid selection events in a herpesvirus, mutations in RL13 and the UL128 locus were also identified at p3 following transfection of the repaired BACs into HFFFs, and these were clearly generated de novo. Interestingly, mutants emerged in RL13 more rapidly than in the UL128 locus. Some data also indicated that RL13 may mutate more rapidly in the context of UL128 mutated viruses; however, this may simply reflect the slower growth of the WT virus.

In support of our conclusions concerning the instability of RL13 in cell culture, a collaborative study has shown that RL13 mutants were invariably selected in fibroblasts, epithelial, and endothelial cells during sequential passage of HCMV strains from clinical samples (28). Also, in most existing HCMV BACs (including FIX, Ph, Towne, Toledo, and AD169), RL13 is mutated overtly by deletions, frameshifts, or substitutions that introduce premature termination codons (27), and that derived from strain TB40/E (54) contains a unique substitution that greatly reduces the prediction of a signal peptide for gpRL13 (A.J. Davison, unpublished observations). Relating to this, we note that on 2 occasions, viruses derived from Merlin BACs replicated relatively efficiently in fibroblasts and yet retained an intact RL13 (RCMV1159 [data not shown] and a proportion of RCMV1161; see Figure 5D). This suggests that another HCMV gene may be required for RL13 function and that this gene is less prone to mutation than RL13.

HCMV exhibits the highest degree of intrastrain sequence variation of any human herpesvirus, and RL13 is a member of a small group of genes that exhibit the greatest variation (27). Also among this group are genes UL146 (27, 35, 36, 66–68) and UL74 (gO) (27, 35, 69–71), which are well characterized because of their utility in genotyping clinical isolates. The selection pressures responsible for generating such degrees of divergence are not fully understood, but their origins, and perhaps the era in which they have operated, appears to be ancient (66). Despite remarkable sequence variation among HCMV strains, UL146 and UL74 are stable within individual patients, and identical UL146 sequences have been detected in geographically distant individuals (35, 66). RL13 is one of 14 members of the RL11 gene family, which is believed to have arisen through gene duplication and then diverged during the evolution of the primate cytomegaloviruses (72). Several other members of this family are also hypervariable. As a virion envelope protein, gpRL13 may be a prime target for neutralizing antibody, and it is possible that selection is exerted on it in vivo to drive escape from the humoral immune response. In support of this, when $CD4^+$ T cell responses to HCMV were measured, RL13 was one of the top 5 most immunogenic HCMV ORFs (73). It would be interesting to determine whether RL13 is now stable or whether the mechanism driving its genetic variation remains operational.

The question arises of why RL13 is detrimental to virus replication in both epithelial and fibroblast cells in culture. The virion components encoded by the UL128 locus are clearly also detrimental in fibroblasts, yet they are essential for infection of other cell types. Since gpRL13 is also present in the virion, it may likewise modify tropism in some way that has not yet been recapitulated in vitro. Recent studies have suggested that HCMV is capable in vivo of establishing persistent, low-level infections that are cor-

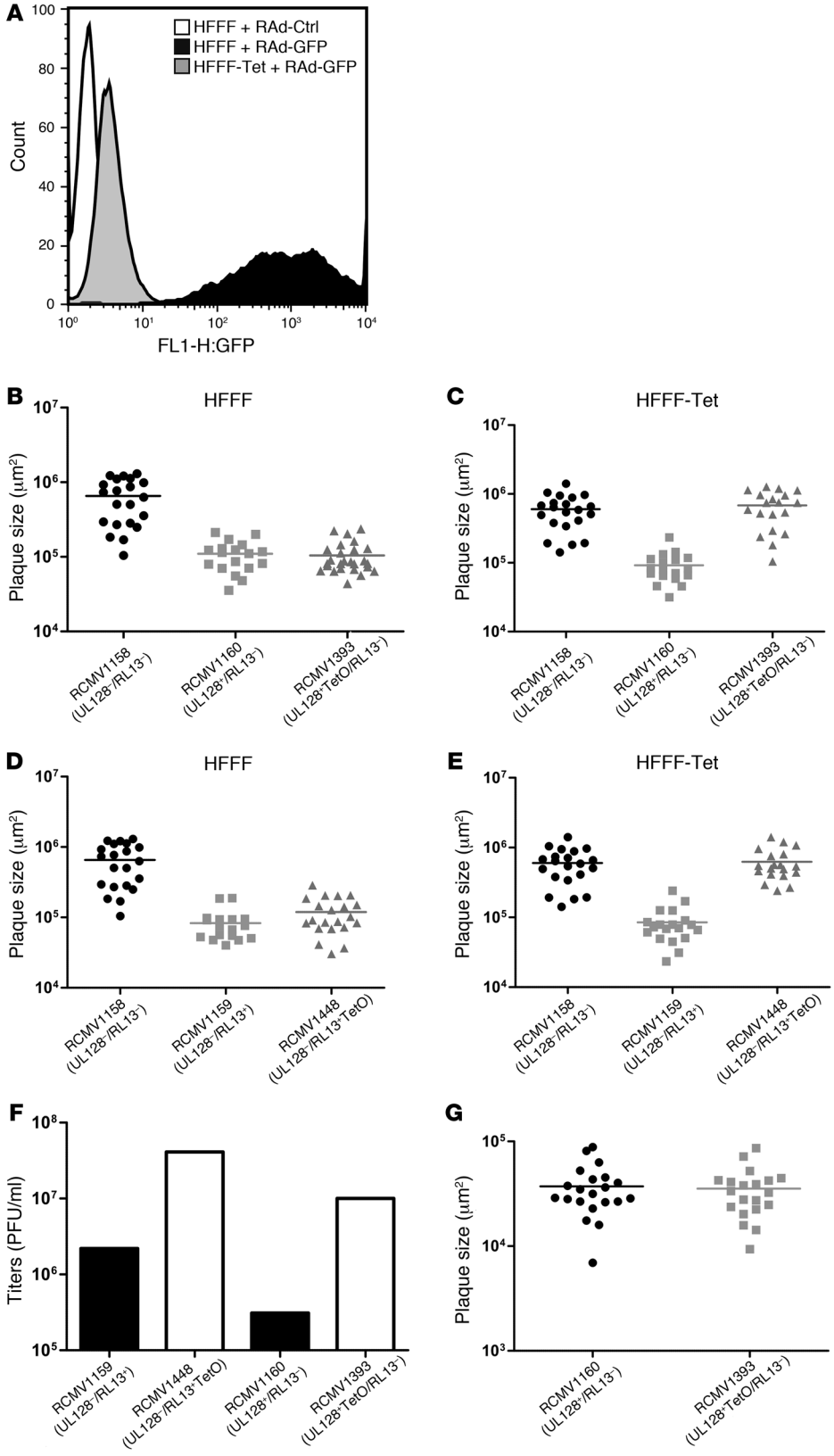


Figure 9
Repression of UL128L and RL13 by BAC-derived viruses in HFFF-tet cells. (**A**) FACS analysis of parental HFFFs or HFFF-tet cells, infected with empty control adenovirus (RAd-Ctrl) or RAd-expressing eGFP (RAd-GFP). RAd-GFP expresses GFP from the HCMV MIE promoter, containing 2 tet operators 10 bp downstream of the TATA box. (**B**–**E**) Plaque sizes of viruses generated at 2 weeks PT in HFFFs (**B** and **D**) or HFFF-tet cells (**C** and **E**), with cells overlaid to prevent cell-free spread of virus, showing (**B** and **C**) repression of UL128L in RCMV1393 and (**D** and **E**) repression of RL13 in RCMV1448. (**F**) Titers of virus stocks obtained from HFFF-tet cells infected with viruses in which RL13 and UL131A were tet controlled (RCMV1448 and RCMV1393, respectively) or the parental viruses in which RL13 and UL131A were not tet controlled (RCMV1159 and RCMV1160, respectively). (**G**) Plaque size of parental virus (RCMV1160) or virus in which UL131A was tet controlled (RCMV1393) in ARPE19 cells, 3 weeks PT.

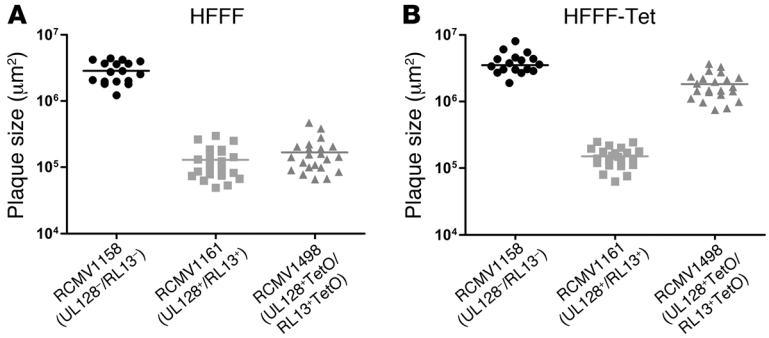


Figure 10
Simultaneous repression of UL128L and UL13 by BAC-derived viruses in HFFF-tet cells. Plaque sizes of viruses generated at 3 weeks PT in HFFFs (**A**) or HFFF-tet (**B**) cells, with cells overlaid to prevent cell-free spread of virus. Data points represent individual plaque sizes recorded for each mutant.

related with the presence of glioblastomas (74–76). Alternatively, therefore, RL13 may act as a regulator, promoting persistence of HCMV by suppressing the switch to full lytic infection until virus dissemination is required. Clearly there are situations in vivo, as illustrated by UL138 (which promotes latency by suppressing lytic infection in $CD34^+$ myeloid progenitors), where it can be beneficial for the virus to restrict or limit productive infection.

Consideration has been given to using live HCMV as a vaccination agent, vaccine carrier, and vector for gene therapy. Such studies are being developed with strains adapted to cell culture that have lost RL13 function. When the low passage HCMV strain Toledo was evaluated in human virus challenge, mild to severe disease was induced in immunocompetent individuals (both seronegative and seropositive) receiving doses ranging between 10 and 1,000 PFU (9). Indeed, the capacity of this strain to induce clinical disease was impressive. However, Toledo is known to have a large, in-frame deletion in RL13 and an ablation of UL128 (27). The absence of RL13 function as a repressor of HCMV replication could potentially have enhanced the efficiency of Toledo replication, dissemination, and pathogenesis during an acute infection. In addition to demonstrating that RL13 is incompatible with efficient virus replication in vitro, we have now provided a method for conditional expression of RL13 and UL131A in the context of productive virus infection in HFFFs. The system that we have developed will facilitate HCMV studies with virus fully representing the clinical agent.

Methods

All studies in human and animal cells were approved by Bro Taf Local Research Ethics Committee, Cardiff, United Kingdom.

Cells and viruses. HFFFs and ARPE-19s (77) were grown in DMEM (Gibco; Invitrogen) containing 10% FCS (Gibco; Invitrogen) at 37°C in 5% CO_2. HCMV strain Merlin (p5) (27) was utilized to generate BACs as described below. Infections were performed as described previously (78), and viruses were titrated in triplicate by plaque assay for 14 days on HFFFs using a 1% Avicel overlay (79). Cultures that exhibited small plaques at 14 days PT were incubated for a further 14 days and recounted. Plaques were visualized using an ORCA-ER camera mounted on a Leica DMIRBE microscope, and sizes were computed using Openlab 3 software (Improvision).

PCR. Three DNA polymerases were used in PCR reactions according to the manufacturers' protocols: Phusion (NEB) for fragments greater than 4 kb; Advantage 2 (Clontech) for amplification directly from virus stocks or cultures; and Expand Hi-Fi (Roche) for all other experiments. Oligonucleotide primers were purchased from Sigma-Aldrich at desalted purity. RL13 and the UL128 locus were amplified for sequencing using primers RL13F (ATCCTGAACATGAAGACTGACGTT) and RL13R (GAATAAACACACCCAAACATTAATGAC), and UL128F (CAGAAACTCACATCGGCGACA) and UL131AR (CCATCACCTCGCCTATACTATGTG), respectively.

Isolation, transformation, and transfection of DNA. Minipreps of plasmids or BACs were produced as described previously (62). Maxipreps were produced using NucleoBond BAC100 kits (Machery Nagel) from 500 ml overnight bacterial cultures. BAC DNA transfections (2 μg) of HFFFs (10^6 cells) were performed using a basic fibroblast kit and a Nucleofector (Amaxa) according to the manufacturer's protocol. Transfections into ARPE-19s were performed in 25-cm^2 flasks using effectene (QIAGEN) according to the manufacturer's protocol for 60-mm dishes.

Transformation of *E. coli* was performed by electroporation at 2.5 kV using 2-mm cuvettes and a Micropulser (Bio-Rad). When PCR products were cloned for sequencing, DNA was purified from agarose gels using a GFX PCR DNA and Gel Band Purification kit (GE Healthcare) and cloned into pCR4-TOPO (Invitrogen) according to the manufacturer's instructions.

Construction of the BAC-targeting vector. The BAC-targeting vector (pAL1026; Figure 1) used to capture the Merlin genome contained markers for eGFP and puromycin resistance (eGFP/Puro) and homology arms matching HCMV genes US28 and US34A (817 and 888 bp, respectively), a chloramphenicol resistance gene (*cat*), genes *sopA*, *sopB*, and *sopC*, which ensure active partitioning during cell division, and gene *repE*, which mediates assembly of the replication complex in *E. coli* at Ori2. Each homology arm was flanked by a *loxP* site, and a unique NheI restriction site was positioned between them.

To generate pAL1026, pBeloBAC11 (New England Biolabs; GenBank U51113; ref. 80) was digested with HpaI and ApaLI, blunt-ended with Klenow DNA polymerase, and religated to remove the existing *cos* and *loxP* sites, thus generating pAL767. Homology arms matching genes US28 and US34 were amplified from Merlin DNA using primers US28F (GGCCGCTAGCTGGCGACGTCGGATTCAATG; NheI site underlined), US28R (GGCCGGATCC**ATAACTTCGTATAATGTATGCTATACGAAGTTAT**AGCGCTTTTTTATTACGGTATAAT; BamHI site underlined, *loxP* site shown in bold), and US34F (GGCCGCATGC**ATAACTTCGTATAGCATACATTATACGAAGTTAT**GACCAGTGGTGGCGGGGAAT; SphI site underlined, *loxP* site shown in bold) and US34R (GGCCGCTAGCGGGACTTTCATCCTGCATTA; NheI site underlined), respectively. The 2 PCR products were digested with NheI, ligated together, digested with SphI and BamHI, and inserted into pAL767, thus generating pAL799. A linker containing a PacI site was inserted into the BamHI site, and then an expression cassette encoding eGFP/Puro (under the control of an SV40 promoter and linked by an internal ribosome entry site [IRES]) was inserted into pAL799, thus generating pAL815. The eGFP/Puro expression cassette was derived from the BAC-targeting vector YD-C19 (a gift from D. Yu; ref. 53) by amplification using primers eGFP-PuroF (GGCCTTAATTAACAATTCGGCGCAGCACCA; PacI site underlined) and eGFP-PuroR (GGCCTTAATTAAGATCCAGACATGATAAGATACATTGATG; PacI site

underlined). The amplified fragment was digested with PacI and inserted into PacI-digested pAL815, thus generating pAL1026.

Construction of a Merlin BAC. 1 μg pAL1026 was linearized by NheI digestion and transfected using Effectene (QIAGEN) into 5×10^5 HFFFs. Cells were infected at 24 hours PT with Merlin (p5) at an MOI of 10, and recombinants were enriched by selection with puromycin (2.5 μg/ml) and detected by visualizing eGFP. When a significant proportion of cells exhibited eGFP expression, circular DNA was extracted using Hirt extraction (53, 81) and transfected into *E. coli*. Selection with chloramphenicol allowed the identification of BAC colonies, from among which pAL1031 (Figure 1) was analyzed in detail and repaired by recombineering.

Repairing of Merlin BACs. Recombineering was performed in *E. coli* SW102 using *lac*Z/*amp*^{r}/*sac*B (62) and *gal*K selection cassettes (82) as previously described. The sequences of the primers used are listed in Supplemental Table 1 (supplemental material available online with this article; doi:10.1172/JCI42955DS1). All constructs were verified by sequencing modified regions and by restriction digest at each step. To enable the Merlin UL29–UL34 sequence to be reintroduced into pAL1031, the *lac*Z/*amp*^{r}/*sac*B cassette was amplified and inserted between the *loxP* site and US34A using primers SacBF-LoxPHom and SacBR-US34AHom to generate pAL1033 (Figure 1). To generate the insert, 2 PCR products were produced, digested with NheI, and ligated together. A 482-bp region encompassing the *loxP* site adjacent to US34 in pAL1026 was amplified using primers LoxP-R (GGCCGCTAGCATAACTTCGTATAATGTATGCTATACGAAGTTATGCATGCAAGCTTGAGTATTC; NheI site underlined) and Chlor-F (GGCCAGATCTGTCGGCAGAATGCTTAATGAA; BglII site underlined). A 5.2-kb region of the Merlin genome spanning genes US29 to US34A was also amplified using primers US29F (GGCCGCTAGCACCTCGGCCTTTTCATACAA; NheI site underlined) and US34AR (GGCCAGATCTGGGACTTTCATCCTGCATTA; BglII site underlined). Each amplicon was cloned into pCR2.1-TOPO (Invitrogen) for sequencing. A clone of each amplicon in which the sequence was correct was digested with BglII/NheI. The fragments were ligated together and recombineered into pAL1033, thus replacing the *lac*Z/*amp*^{r}/*sac*B cassette by US29–US34A and generating pAL1040 (Figure 1).

The eGFP/Puro cassette in pAL1040 was replaced by a Cre recombinase gene under the control of an SV40 promoter and containing a synthetic intron (57) so that the protein would be expressed in mammalian cells but not *E. coli*. The *lac*Z/*amp*^{r}/*sac*B cassette was amplified using primers SacBF-GFPPuroHom and SacBR-GFPPuroHom and inserted into pAL1040, thus generating pAL1047 (Figure 1). The Cre recombinase expression cassette was amplified from YD-C66 (a gift from D. Yu; ref. 53), using primers CreF-SV40Hom and CreR-BACHom, and recombineered in place of the *lac*Z/*amp*^{r}/*sac*B cassette, thus generating pAL1053 (Figure 1).

The sequence of pAL1053 was determined (see below) and indicated the presence of several differences from that of WT Merlin (see Results). Redundant sequences containing residual *lac*Z sequences were deleted by replacement with a *gal*K selectable marker, which was amplified using primers GalKF-CreHom and GalKR-SopCHom. The *gal*K marker was then removed using the primer RemoveGalKBAC, thus generating pAL1090 (Figure 1). To repair the lesions in gene UL36, the *lac*Z/*amp*^{r}/*sac*B cassette was amplified using primers SacBF-UL36Hom and SacBR-UL36Hom and inserted in place of UL36. UL36 was then amplified from Merlin DNA using primers UL36F and UL36R and recombineered in place of the *lac*Z/*amp*^{r}/*sac*B cassette, thus generating pAL1111 (Figure 1). Additional BACs were generated by repairing either and then both of the UL128 and RL13 mutations by the same technique. In each case, the *lac*Z/*amp*^{r}/*sac*B cassette was amplified using primers flanking the mutation (SacBF-UL128Hom and SacBR-UL128Hom or SacBF-RL13Hom and SacBR-RL13Hom) and recombineered into the relevant BAC, and the selectable marker was excised and the mutation repaired by recombination with oligonucleotide UL128Rep or RL13Rep.

To insert an IRES-eGFP expression cassette immediately downstream from gene UL122 (IE2) into each of the 4 BACs described above, the *lac*Z/*amp*^{r}/*sac*B cassette was amplified using primers SacBF-IE2Hom and SacBR-IE2Hom. In the resulting BACs, the *lac*Z/*amp*^{r}/*sac*B cassette was replaced by an IRES-eGFP cassette amplified from pIRES2-GFP (Clontech) using primers IRESF-IE2Hom and GFPR-IE2Hom.

DNA sequencing. All PCR products in directly cloned or recombineered form were verified by sequencing using a BigDye 3.1 kit (ABI) and standard techniques, except that the number of cycles during sequencing was increased from 25 to 100. Reactions were purified using Dye Terminator Removal Columns (EdgeBio) and analyzed on an ABI Prism 3130xl Genetic Analyzer (Applied Biosystems).

Two BACs were sequenced using published approaches (59). The complete HCMV component of pAL1053 was determined via standard Staden sequencing of a set of overlapping PCR products, and that of pAL1128 in its entirety was determined using data from an Illumina Genome Analyzer (The GenePool, University of Edinburgh). The sequence of pAL1128 was deposited in GenBank as accession GU179001. All nts in the text are specified relative to Merlin (NC_006273).

QRT-PCR. QRT-PCR was performed to analyze the expression levels of gene US29 in infected cells. The cells were trypsinized and washed once in PBS, and whole-cell RNA was extracted using an RNEasy Mini Kit (QIAGEN) and DNAse-treated using Turbo DNA-free (Ambion). US29 transcripts were amplified using primers US29F (GACATCGGTGACACAGCTTCA) and US29R (CTTGGCCTTCAGAGACCGC), and UL123 (IE1) transcripts were amplified using primers IE1F (AGGAAGAAAGTGAACAGAGTGATGA) and IE1R (TTCCTCAGCACCATCCTCCT). For RT-PCR with real-time detection, an iScript One-Step RT-PCR kit with SYBR Green (Bio-Rad) was employed according to manufacturer's instructions, and samples were analyzed on an iCycler (Bio-Rad). Serial dilutions of a Merlin BAC (pAL1053; see below) were used to establish a calibration curve. Since transcripts antisense to US29 have been reported (83), the sense transcript was amplified specifically using the reverse primers (100 pmol) with 100 ng RNA in a reverse transcription step at 50°C for 10 minutes, followed by inactivating reverse transcriptase at 95°C for 5 minutes. The forward primer was then added, and the samples were cycled 40 times through 95°C for 10 seconds, 60°C 30 seconds, and 83°C for 10 seconds, with data being collected following the final step of each cycle. Melt curve analysis was performed following the 40 cycles, and amplification of a product of the correct size was confirmed by agarose gel electrophoresis. Samples were analyzed in triplicate, and the results were averaged. Samples from uninfected cells and infected cells processed without conducting the reverse transcriptase step were uniformly negative.

Flow cytometry. A Cytofix/Cytoperm plus kit (BD) was used for intracellular staining of the US28 protein, and the Tub-45 antibody (a gift from U. Höpken; ref. 84) was used (1:5) in combination with goat anti-mouse Alexa Fluor 488 (1:200) for detection (Invitrogen). HCMV infection was monitored directly by eGFP expression or indirectly by downregulation of MHC class I from the cell surface using antibody W6/32 (1:100) followed by goat anti-mouse Alexa Fluor 647 (1:200) (Invitrogen). Samples were measured on a FACSCalibur (BD), and data were analyzed in FlowJo (Treestar).

Generation of the HFFF-tet cell line. The coding region of the tet repressor was amplified using primers tetR-F (GGCCGGATCCCTATAGGGCGAATTGAT**ATG**TCTA; BamHI site underlined, initiation codon in bold) and tetR-R (GGCCGAATTC**TTATACCTTTCTCTTCTTTTTTGGATCA**GACCCACTTTCACATTTAAGTTG; EcoRI site underlined, NLS shown in bold) and cloned into the BamHI/EcoRI sites of the retrovirus vector pMXs-IP (85). Retrovirus stocks were produced by transfecting 293Phoenix

packaging cells (86) using Effectene (QIAGEN) according to manufacturer's instructions and harvesting the supernatant 48 hours PT. Retrovirus was bound to culture dishes with Retronectin (Clontech) and used to infect hTERT-immortalized HFFFs. The cells were selected in puromycin (1 μg/ml) at 48 hours PI.

Inserting tet operators into BACs. tet operators were inserted by recombineering. To insert 1 tet operator 19 bp upstream of the RL13 coding region, the *lacZ*/*amp*^{r}/*sac*B cassette was amplified with primers SacBF-RL13-2 and SacBR-RL13-2 and inserted into the BAC and then removed with oligonucleotide tetO-RL13, leaving behind 1 tet operator. To insert 2 tet operators 33 bp upstream of the UL131A coding region, the *lacZ*/*amp*^{r}/*sac*B cassette was amplified with primers SacBF-UL131A and SacBR-UL131A and inserted into the BAC and then removed with oligonucleotide tetO-UL131A, leaving behind 2 tet operators.

Tagging RL13 in BACs with a V5 epitope. A sequence encoding a V5 epitope was fused to the C terminus of gpRL13 by recombineering. The *lacZ*/*amp*^{r}/*sac*B cassette was amplified with primers SacBF-RL13Hom and SacBR-RL13Hom and inserted after the RL13 ORF within the BAC; then the cassette was removed using overlapping oligos V5-RL13-F and V5-RL13-R, which fused the coding sequence for a V5 tag to the end of gene RL13.

Generating replication-deficient RAd vectors. RL13 was cloned by recombineering into a RAd using the AdZ system as described previously (62), using primers RL13F and RL13R, so that the gpRL13 was tagged with a C-terminal V5 epitope. Virus was recovered from the vector by transfection into 293TREx cells and titrated as described previously (62). Empty vector control (RAd-Ctrl) and RAd-GFP, which expresses eGFP from a version of the HCMV MIE promoter that contains 2 tet operators 10 bp downstream from the TATA box, have been described previously (62).

Protein predictions. N- and O-linked glycosylation sites were predicted using NetNGlyc 1.0 and NetOGlyc 3.1 (87). Signal sequences and transmembrane domains were predicted using SignalP 3.0 (88) and TMHMM 2.0 (http://www.cbs.dtu.dk/services/).

Purification of HCMV virions. HCMV particles in cell supernatants were separated into virion, dense body, and noninfectious enveloped particle (NIEP) fractions by positive density/negative viscosity gradient centrifugation as described previously (89). Particle concentrations in the preparations were estimated by counting negatively stained samples by EM in relation to a standard concentration of latex beads. To separate virion envelope proteins from capsid and tegument proteins, 10^8 particles were mixed 1:1 with envelope stripping buffer (2% Nonidet-P40 in PBS) and incubated for 15 minutes at 4°C. Particles were pelleted (12,000 *g* for 5 minutes at 4°C), and the soluble envelope fraction was harvested. The insoluble capsid/tegument material was washed twice with envelope-stripping buffer and once in PBS before being solubilized in SDS-PAGE sample buffer.

SDS-PAGE and Western immunoblotting. Protein samples were separated by SDS-PAGE using 4%–12% Bis-Tris NuPAGE protein gels (Invitrogen) according to the manufacturer's instructions, then transferred to nitrocellulose by semi-dry transfer. Membranes were blocked overnight in blocking buffer (3% BSA in PBS containing 0.1% Tween 20 [PBST]) and then incubated with primary antibody for 1 hour. Following 3 washes in PBST, secondary antibody was incubated for 1 hour and washed 3 times. Bound antibody was detected using ECL–Western blotting detection system (RPN 2132; Amersham) and exposure to film. Primary antibodies were chicken anti-V5 (9113; Abcam), mouse anti-gB (CMV-023-40154; Capricorn), and mouse anti-pp65 (CMV-018-48151; Capricorn). Secondary antibodies were goat anti–chicken-HRP (6877; Abcam) and goat anti–mouse-HRP (170-6516; Bio-Rad).

Glycosidase treatment. Samples were treated with PNGase F (P0705S; New England BioLabs) or ENDO H (P0703S; New England BioLabs) according to the manufacturer's instructions. Briefly, the samples were denatured in glycoprotein-denaturing buffer at 100°C for 10 minutes and cooled to 0°C for 5 minutes. The samples were then digested overnight at 37°C with PNGase F or ENDO H before being analyzed by Western immunoblotting.

Immunogold electron microscopy. Purified virions were air-dried to the surface of Formvar-coated EM grids. The grids were treated with chicken anti-V5 antibody (9113) for 4 hours at room temperature, washed 3 times with PBS, and incubated with goat anti-chicken antibody (39604; Abcam) conjugated to 6-nm gold particles for 1 hour at room temperature. After further washing in PBS, the grids were negatively stained with phosphotungstic acid and subjected to EM.

Immunofluorescence. Cells were plated on glass coverslips and infected with HCMV. At 7 days PI, the cells were fixed in 4% paraformaldehyde, permeabilized with 0.5% NP-40, and incubated with primary antibody for 1 hour at 37°C. Following washing, secondary antibodies were incubated for 1 hour at 37°C, washed again, and mounted. Primary antibodies were chicken anti-V5 (9113; Abcam), sheep anti-human TGN46 (AHP500; Serotec), mouse MAb anti-pp28 (6502; Abcam), and mouse anti-gH (2470-5437; Biogenesis). Secondary antibodies were donkey anti-chicken Cy3 (AP194C; Millipore), donkey anti-sheep IgG–Alexa Fluor 633 (A21100; Invitrogen), and donkey anti-mouse IgG–Alexa Fluor 488 (A-2102; Invitrogen). The intracellular locations of antibody-tagged proteins were examined under laser illumination in a Zeiss LMS 510 confocal microscope, and images were captured using LMS software.

Statistics. Means ± SD were calculated using GraphPad Prims software.

Acknowledgments

This work was supported by the Wellcome Trust, the United Kingdom Medical Research Council (MRC), and the Biotechnology and Biological Sciences Research Council (BBSRC). The authors thank Sian Llewellyn-Lacey and Carole Rickards for tissue culture support; Brent Ryckman for helpful discussions; Neal Copeland, Dong Yu, and Uta Höpken for reagents; Richard Darley and Phillip Taylor for reagents and advice regarding retrovirus generation; and Jim Aitken for performing the EM.

Received for publication March 12, 2010, and accepted in revised form June 23, 2010.

Address correspondence to: Richard Stanton, Section of Medical Microbiology, Department of Infection, Immunity and Biochemistry, School of Medicine, Cardiff University, Tenovus Building, Heath Park, Cardiff CF14 4XN, United Kingdom. Phone: 44.0.29.20687319; Fax: 44.0.29.20742161; E-mail: StantonRJ@cf.ac.uk.

1. Mocarski ES. Cytomegaloviruses. In: Fields BN, Knipe DM, Howley PM, eds. *Fields Virology*. 5th ed. Philadelphia, Pennsylvania, USA: Lippincott–Raven; 2006:2447–2492.
2. Stratton KR, Durch JS, Lawrence RS. *Vaccines for the 21st Century: A Tool for Decisionmaking*. Washington, DC, USA: National Academies Press; 2000.
3. Sinzger C, Grefte A, Plachter B, Gouw AS, The TH, Jahn G. Fibroblasts, epithelial cells, endothelial cells and smooth muscle cells are major targets of human cytomegalovirus infection in lung and gastrointestinal tissues. *J Gen Virol*. 1995;76(pt 4):741–750.
4. Sinzger C, Plachter B, Grefte A, The TH, Jahn G. Tissue macrophages are infected by human cytomegalovirus in vivo. *J Infect Dis*. 1996;173(1):240–245.
5. Bissinger AL, Sinzger C, Kaiserling E, Jahn G. Human cytomegalovirus as a direct pathogen: correlation of multiorgan involvement and cell distribution with clinical and pathological findings in a case of congenital inclusion disease. *J Med Virol*. 2002;67(2):200–206.
6. Kahl M, Siegel-Axel D, Stenglein S, Jahn G, Sinzger C. Efficient lytic infection of human arterial endothelial cells by human cytomegalovirus strains. *J Virol*. 2000;74(16):7628–7635.
7. Riegler S, Hebart H, Einsele H, Brossart P, Jahn G, Sinzger C. Monocyte-derived dendritic cells are permissive to the complete replicative cycle of human cytomegalovirus. *J Gen Virol*. 2000;81(pt 2):393–399.
8. Quinnan GV Jr, et al. Comparative virulence and

immunogenicity of the Towne strain and a nonattenuated strain of cytomegalovirus. *Ann Intern Med.* 1984;101(4):478–483.
9. Plotkin SA, Starr SE, Friedman HM, Gonczol E, Weibel RE. Protective effects of Towne cytomegalovirus vaccine against low-passage cytomegalovirus administered as a challenge. *J Infect Dis.* 1989;159(5):860–865.
10. Elek SD, Stern H. Development of a vaccine against mental retardation caused by cytomegalovirus infection in utero. *Lancet.* 1974;1(7845):1–5.
11. Cha TA, Tom E, Kemble GW, Duke GM, Mocarski ES, Spaete RR. Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains. *J Virol.* 1996;70(1):78–83.
12. Prichard MN, Penfold ME, Duke GM, Spaete RR, Kemble GW. A review of genetic differences between limited and extensively passaged human cytomegalovirus strains. *Rev Med Virol.* 2001;11(3):191–200.
13. Davison AJ, et al. The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. *J Gen Virol.* 2003; 84(pt 1):17–28.
14. Skaletskaya A, Bartle LM, Chittenden T, McCormick AL, Mocarski ES, Goldmacher VS. A cytomegalovirus-encoded inhibitor of apoptosis that suppresses caspase-8 activation. *Proc Natl Acad Sci U S A.* 2001;98(14):7829–7834.
15. Dargan DJ, Jamieson FE, MacLean J, Dolan A, Addison C, McGeoch DJ. The published DNA sequence of human cytomegalovirus strain AD169 lacks 929 base pairs affecting genes UL42 and UL43. *J Virol.* 1997;71(12):9833–9836.
16. Mocarski ES, Prichard MN, Tan CS, Brown JM. Reassessing the organization of the UL42-UL43 region of the human cytomegalovirus strain AD169 genome. *Virology.* 1997;239(1):169–175.
17. Brown JM, Kaneshima H, Mocarski ES. Dramatic interstrain differences in the replication of human cytomegalovirus in SCID-hu mice. *J Infect Dis.* 1995;171(6):1599–1603.
18. Bradley AJ, et al. High-throughput sequence analysis of variants of human cytomegalovirus strains Towne and AD169. *J Gen Virol.* 2009; 90(pt 10):2375–2380.
19. Penfold ME, et al. Cytomegalovirus encodes a potent alpha chemokine. *Proc Natl Acad Sci U S A.* 1999;96(17):9839–9844.
20. Benedict CA, et al. Cutting edge: a novel viral TNF receptor superfamily member in virulent strains of human cytomegalovirus. *J Immunol.* 1999; 162(12):6967–6970.
21. Wills MR, et al. Human cytomegalovirus encodes an MHC class I-like molecule (UL142) that functions to inhibit NK cell lysis. *J Immunol.* 2005; 175(11):7457–7465.
22. Tomasec P, et al. Downregulation of natural killer cell-activating ligand CD155 by human cytomegalovirus UL141. *Nat Immunol.* 2005;6(2):181–188.
23. Cerboni C, et al. Human cytomegalovirus strain-dependent changes in NK cell recognition of infected fibroblasts. *J Immunol.* 2000;164(9):4775–4782.
24. Goodrum F, Reeves M, Sinclair J, High K, Shenk T. Human cytomegalovirus sequences expressed in latently infected individuals promote a latent infection in vitro. *Blood.* 2007;110(3):937–945.
25. Petrucelli A, Rak M, Grainger L, Goodrum F. Characterization of a novel Golgi apparatus-localized latency determinant encoded by human cytomegalovirus. *J Virol.* 2009;83(11):5615–5629.
26. Akter P, et al. Two novel spliced genes in human cytomegalovirus. *J Gen Virol.* 2003;84(pt 5):1117–1122.
27. Dolan A, et al. Genetic content of wild-type human cytomegalovirus. *J Gen Virol.* 2004; 85(pt 5):1301–1312.
28. Dargan DJ, et al. Sequential mutations associated with adaptation of human cytomegalovirus to growth in cell culture. *J Gen Virol.* 2010;91(pt 6):1535–1546.
29. Chandler SH, Handsfield HH, McDougall JK. Isolation of multiple strains of cytomegalovirus from women attending a clinic for sexually transmitted disease. *J Infect Dis.* 1987;155(4):655–660.
30. Collier AC, Chandler SH, Handsfield HH, Corey L, McDougall JK. Identification of multiple strains of cytomegalovirus in homosexual men. *J Infect Dis.* 1989;159(1):123–126.
31. Spector SA, Hirata KK, Newman TR. Identification of multiple cytomegalovirus strains in homosexual men with acquired immunodeficiency syndrome. *J Infect Dis.* 1984;150(6):953–956.
32. Baldanti F, et al. Coinfection of the immunocompromised but not the immunocompetent host by multiple human cytomegalovirus strains. *Arch Virol.* 1998;143(9):1701–1709.
33. Chou SW. Reactivation and recombination of multiple cytomegalovirus strains from individual organ donors. *J Infect Dis.* 1989;160(1):11–15.
34. Drew WL, Sweet ES, Miner RC, Mocarski ES. Multiple infections by cytomegalovirus in patients with acquired immunodeficiency syndrome: documentation by Southern blot hybridization. *J Infect Dis.* 1984;150(6):952–953.
35. Stanton R, Westmoreland D, Fox JD, Davison AJ, Wilkinson GW. Stability of human cytomegalovirus genotypes in persistently infected renal transplant recipients. *J Med Virol.* 2005;75(1):42–46.
36. Hassan-Walker AF, Okwuadi S, Lee L, Griffiths PD, Emery VC. Sequence variability of the alpha-chemokine UL146 from clinical strains of human cytomegalovirus. *J Med Virol.* 2004;74(4):573–579.
37. Adler B, Scrivano L, Ruzcics Z, Rupp B, Sinzger C, Koszinowski U. Role of human cytomegalovirus UL131A in cell type-specific virus entry and release. *J Gen Virol.* 2006;87(pt 9):2451–2460.
38. Hahn G, et al. Human cytomegalovirus UL131-128 genes are indispensable for virus growth in endothelial cells and virus transfer to leukocytes. *J Virol.* 2004;78(18):10023–10033.
39. Patrone M, Secchi M, Bonaparte E, Milanesi G, Gallina A. Cytomegalovirus UL131-128 products promote gB conformational transition and gB-gH interaction during entry into endothelial cells. *J Virol.* 2007;81(20):11479–11488.
40. Ryckman BJ, Chase MC, Johnson DC. HCMV gH/gL/UL128-131 interferes with virus entry into epithelial cells: Evidence for cell type-specific receptors. *Proc Natl Acad Sci U S A.* 2008;105(37):14118–14123.
41. Ryckman BJ, Jarvis MA, Drummond DD, Nelson JA, Johnson DC. Human cytomegalovirus entry into epithelial and endothelial cells depends on genes UL128 to UL150 and occurs by endocytosis and low-pH fusion. *J Virol.* 2006;80(2):710–722.
42. Ryckman BJ, et al. Characterization of the human cytomegalovirus gH/gL/UL128-131 complex that mediates entry into epithelial and endothelial cells. *J Virol.* 2008;82(1):60–70.
43. Wang D, Shenk T. Human cytomegalovirus UL131 open reading frame is required for epithelial cell tropism. *J Virol.* 2005;79(16):10330–10338.
44. Waldman WJ, Roberts WH, Davis DH, Williams MV, Sedmak DD, Stephens RE. Preservation of natural endothelial cytopathogenicity of cytomegalovirus by propagation in endothelial cells. *Arch Virol.* 1991;117(3–4):143–164.
45. Sinzger C, et al. Modification of human cytomegalovirus tropism through propagation in vitro is associated with changes in the viral genome. *J Gen Virol.* 1999;80(pt 11):2867–2877.
46. Sinzger C, Knapp J, Plachter B, Schmidt K, Jahn G. Quantification of replication of clinical cytomegalovirus isolates in cultured endothelial cells and fibroblasts by a focus expansion assay. *J Virol Methods.* 1997;63(1–2):103–112.
47. MacCormac LP, Grundy JE. Two clinical isolates and the Toledo strain of cytomegalovirus contain endothelial cell tropic variants that are not present in the AD169, Towne, or Davis strains. *J Med Virol.* 1999;57(3):298–307.
48. Grazia Revello M, et al. In vitro selection of human cytomegalovirus variants unable to transfer virus and virus products from infected cells to polymorphonuclear leukocytes and to grow in endothelial cells. *J Gen Virol.* 2001;82(pt 6):1429–1438.
49. Marchini A, Liu H, Zhu H. Human cytomegalovirus with IE-2 (UL122) deleted fails to express early lytic genes. *J Virol.* 2001;75(4):1870–1878.
50. Murphy E, et al. Coding potential of laboratory and clinical strains of human cytomegalovirus. *Proc Natl Acad Sci U S A.* 2003;100(25):14976–14981.
51. Hahn G, Rose D, Wagner M, Rhiel S, McVoy MA. Cloning of the genomes of human cytomegalovirus strains Toledo, TownevarRIT3, and Towne long as BACs and site-directed mutagenesis using a PCR-based technique. *Virology.* 2003;307(1):164–177.
52. Borst EM, Hahn G, Koszinowski UH, Messerle M. Cloning of the human cytomegalovirus (HCMV) genome as an infectious bacterial artificial chromosome in Escherichia coli: a new approach for construction of HCMV mutants. *J Virol.* 1999; 73(10):8320–8329.
53. Yu D, Smith GA, Enquist LW, Shenk T. Construction of a self-excisable bacterial artificial chromosome containing the human cytomegalovirus genome and mutagenesis of the diploid TRL/IRL13 gene. *J Virol.* 2002;76(5):2316–2328.
54. Sinzger C, et al. Cloning and sequencing of a highly productive, endotheliotropic virus strain derived from human cytomegalovirus TB40/E. *J Gen Virol.* 2008;89(pt 2):359–368.
55. Copeland NG, Jenkins NA, Court DL. Recombineering: a powerful new tool for mouse functional genomics. *Nat Rev Genet.* 2001;2(10):769–779.
56. Court DL, Sawitzke JA, Thomason LC. Genetic engineering using homologous recombination. *Annu Rev Genet.* 2002;36:361–388.
57. Smith GA, Enquist LW. A self-recombining bacterial artificial chromosome and its application for analysis of herpesvirus pathogenesis. *Proc Natl Acad Sci U S A.* 2000;97(9):4873–4878.
58. Hobom U, Brune W, Messerle M, Hahn G, Koszinowski UH. Fast screening procedures for random transposon libraries of cloned herpesvirus genomes: mutational analysis of human cytomegalovirus envelope glycoprotein genes. *J Virol.* 2000; 74(17):7720–7729.
59. Cunningham C, et al. Sequences of complete human cytomegalovirus genomes from infected cell cultures and clinical specimens. *J Gen Virol.* 2010; 91(pt 3):605–615.
60. Sanchez V, Clark CL, Yen JY, Dwarakanath R, Spector DH. Viable human cytomegalovirus recombinant virus with an internal deletion of the IE2 86 gene affects late stages of viral replication. *J Virol.* 2002;76(6):2973–2989.
61. Griffin C, et al. Characterization of a highly glycosylated form of the human cytomegalovirus HLA class I homologue gpUL18. *J Gen Virol.* 2005; 86(pt 11):2999–3008.
62. Stanton RJ, McSharry BP, Armstrong M, Tomasec P, Wilkinson GW. Re-engineering adenovirus vector systems to enable high-throughput analyses of gene function. *Biotechniques.* 2008;45(6):659–662, 664–668.
63. Yao F, Svensjo T, Winkler T, Lu M, Eriksson C, Eriksson E. Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. *Hum Gene Ther.* 1998;9(13):1939–1950.
64. Lilja AE, Shenk T. Efficient replication of rhesus cytomegalovirus variants in multiple rhesus and human cell types. *Proc Natl Acad Sci U S A.* 2008;105(50):19950–19955.
65. Drake JW, Hwang CB. On the mutation rate of herpes simplex virus type 1. *Genetics.* 2005;170(2):969–970.

66. Bradley AJ, et al. Genotypic analysis of two hypervariable human cytomegalovirus genes. *J Med Virol*. 2008;80(9):1615–1623.
67. Lurain NS, et al. Analysis of the human cytomegalovirus genomic region from UL146 through UL147A reveals sequence hypervariability, genotypic stability, and overlapping transcripts. *Virol J*. 2006;3:4.
68. He R, et al. Sequence variability of human cytomegalovirus UL146 and UL147 genes in low-passage clinical isolates. *Intervirology*. 2006;49(4):215–223.
69. Mattick C, et al. Linkage of human cytomegalovirus glycoprotein gO variant groups identified from worldwide clinical isolates with gN genotypes, implications for disease associations and evidence for N-terminal sites of positive selection. *Virology*. 2004; 318(2):582–597.
70. Rasmussen L, Geissler A, Cowan C, Chase A, Winters M. The genes encoding the gCIII complex of human cytomegalovirus exist in highly diverse combinations in clinical isolates. *J Virol*. 2002;76(21):10841–10848.
71. Rasmussen L, Geissler A, Winters M. Inter- and intragenic variations complicate the molecular epidemiology of human cytomegalovirus. *J Infect Dis*. 2003;187(5):809–819.
72. Davison AJ, et al. Homology between the human cytomegalovirus RL11 gene family and human adenovirus E3 genes. *J Gen Virol*. 2003;84(pt 3):657–663.
73. Sylwester AW, et al. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. *J Exp Med*. 2005;202(5):673–685.
74. Cobbs CS, et al. Human cytomegalovirus infection and expression in human malignant glioma. *Cancer Res*. 2002;62(12):3347–3350.
75. Luo MH, Fortunato EA. Long-term infection and shedding of human cytomegalovirus in T98G glioblastoma cells. *J Virol*. 2007;81(19):10424–10436.
76. Mitchell DA, et al. Sensitive detection of human cytomegalovirus in tumors and peripheral blood of patients diagnosed with glioblastoma. *Neuro Oncol*. 2008;10(1):10–18.
77. Dunn KC, Aotaki-Keen AE, Putkey FR, Hjelmeland LM. ARPE-19, a human retinal pigment epithelial cell line with differentiated properties. *Exp Eye Res*. 1996; 62(2):155–169.
78. Stanton RJ, McSharry BP, Rickards CR, Wang EC, Tomasec P, Wilkinson GW. Cytomegalovirus destruction of focal adhesions revealed in a high throughput Western blot analysis of cellular protein expression. *J Virol*. 2007;81(15):7860–7872.
79. Matrosovich M, Matrosovich T, Garten W, Klenk HD. New low-viscosity overlay medium for viral plaque assays. *Virol J*. 2006;3:63.
80. Shizuya H, et al. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proc Natl Acad Sci U S A*. 1992;89(18):8794–8797.
81. Hirt B. Selective extraction of polyoma DNA from infected mouse cell cultures. *J Mol Biol*. 1967; 26(2):365–369.
82. Warming S, Costantino N, Court DL, Jenkins NA, Copeland NG. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res*. 2005;33(4):e36.
83. Grey F, et al. Identification and characterization of human cytomegalovirus-encoded microRNAs. *J Virol*. 2005;79(18):12095–12099.
84. Mokros T, Rehm A, Droese J, Oppermann M, Lipp M, Hopken UE. Surface expression and endocytosis of the human cytomegalovirus-encoded chemokine receptor US28 is regulated by agonist-independent phosphorylation. *J Biol Chem*. 2002;277(47):45122–45128.
85. Kitamura T, et al. Retrovirus-mediated gene transfer and expression cloning: powerful tools in functional genomics. *Exp Hematol*. 2003;31(11):1007–1014.
86. Kinsella TM, Nolan GP. Episomal vectors rapidly and stably produce high-titer recombinant retrovirus. *Hum Gene Ther*. 1996;7(12):1405–1413.
87. Julenius K, Molgaard A, Gupta R, Brunak S. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology*. 2005;15(2):153–164.
88. Emanuelsson O, Brunak S, von Heijne G, Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*. 2007;2(4):953–971.
89. Irmiere A, Gibson W. Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. *Virology*. 1983;130(1):118–133.

High-resolution human cytomegalovirus transcriptome

Derek Gatherer[a], Sepehr Seirafian[b], Charles Cunningham[a], Mary Holton[a], Derrick J. Dargan[a], Katarina Baluchova[a,1], Ralph D. Hector[a,2], Julie Galbraith[c], Pawel Herzyk[c], Gavin W. G. Wilkinson[b], and Andrew J. Davison[a,3]

[a]Medical Research Council–University of Glasgow Centre for Virus Research, Glasgow G11 5JR, United Kingdom; [b]School of Medicine, Cardiff University, Cardiff CF14 4XN, United Kingdom; and [c]Sir Henry Wellcome Functional Genomics Facility, Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, United Kingdom

Edited by Elliott Kieff, Harvard Medical School and Brigham and Women's Hospital, Boston, MA, and approved October 27, 2011 (received for review September 26, 2011)

Deep sequencing was used to bring high resolution to the human cytomegalovirus (HCMV) transcriptome at the stage when infectious virion production is under way, and major findings were confirmed by extensive experimentation using conventional techniques. The majority (65.1%) of polyadenylated viral RNA transcription is committed to producing four noncoding transcripts (RNA2.7, RNA1.2, RNA4.9, and RNA5.0) that do not substantially overlap designated protein-coding regions. Additional noncoding RNAs that are transcribed antisense to protein-coding regions map throughout the genome and account for 8.7% of transcription from these regions. RNA splicing is more common than recognized previously, which was evidenced by the identification of 229 potential donor and 132 acceptor sites, and it affects 58 protein-coding genes. The great majority (94) of 96 splice junctions most abundantly represented in the deep-sequencing data was confirmed by RT-PCR or RACE or supported by involvement in alternative splicing. Alternative splicing is frequent and particularly evident in four genes (*RL8A*, *UL74A*, *UL124*, and *UL150A*) that are transcribed by splicing from any one of many upstream exons. The analysis also resulted in the annotation of four previously unrecognized protein-coding regions (*RL8A*, *RL9A*, *UL150A*, and *US33A*), and expression of the *UL150A* protein was shown in the context of HCMV infection. The overall conclusion, that HCMV transcription is complex and multifaceted, has implications for the potential sophistication of virus functionality during infection. The study also illustrates the key contribution that deep sequencing can make to the genomics of nuclear DNA viruses.

The genetic repertoire of human cytomegalovirus (HCMV; species *Human herpesvirus 5*) is incompletely understood. Most bioinformatic investigations have focused on identifying open reading frames (ORFs) that are conserved in other organisms or that exhibit pattern-based similarities (e.g., in nucleotide or codon bias) to recognized protein-coding regions (CRs) (1). Our current map of the wild-type HCMV genome, based on strain Merlin, contains 166 protein-coding genes (2–5). It is entirely possible that additional, small protein-coding genes will be found. Candidates involve ORFs that overlap recognized CRs and for which there is some evidence for expression (6), ORFs highlighted in pattern-based bioinformatic (7) or proteomic analyses (8), and ORFs whose expression is presently unsuspected.

Recognition of many protein-coding genes has been supplemented by information on protein expression and function. However, HCMV also specifies polyadenylated (polyA) transcripts that, because they lack sizeable, conserved ORFs, appear unlikely to function via translation. One class consists of noncoding, nonoverlapping transcripts (NNTs) that do not substantially overlap the designated CRs of other genes. These include an abundant 2.7-kb RNA (β2.7 or RNA2.7) (9), a 1.1-kb spliced RNA and associated 5-kb stable intron (RNA5.0) (10, 11), and a 1.2-kb RNA (RNA1.2) (12). RNA2.7 functions in inhibiting apoptosis (13), the RNA5.0 homolog in murine cytomegalovirus is involved in virulence (14), and the role of RNA1.2 is unknown. Antisense transcripts (ASTs) form a second class of noncoding RNA, and are transcribed antisense to CRs (15, 16). In addition, HCMV specifies several small, nonpolyA RNAs that may have regulatory roles, including microRNAs (miRNAs) (17) processed from longer transcripts and RNAs associated with the origin of DNA replication (18).

To extend the understanding of HCMV genome expression, we used deep sequencing to study viral transcription in human fibroblasts. The focus was on identifying previously unrecognized protein-coding or noncoding polyA RNAs, and on determining the extent of RNA splicing.

Results

Transcription Profiles. Whole-cell RNA was isolated from human fibroblasts infected with HCMV strain Merlin at 72 h after infection, when infectious virion production was under way. An Illumina DNA sequence dataset (DS) was generated from each of two RNA preparations: a nondirectional DS (NDS; 24,439,691 reads) and a directional DS (DDS; 32,514,724 reads). Transcript orientation could be inferred from the DDS but not the NDS.

The DSs were assembled against the 236-kbp HCMV genome, which has the structure *ab*-U_L-*b′a′c′*-U_S-*ca*, where U_L and U_S are unique regions flanked by inverted repeats (*ab*/*b′a′* and *a′c′*/*ca*). Similar percentages (42.24 and 43.03, respectively) of NDS and DDS reads originated from HCMV. The light yellow windows in Fig. 1 depict transcription profiles, calculated from read densities across the genome. In the lower windows, the NDS profile (gray) on a linear scale demonstrates that some regions are transcribed abundantly (a few off-scale), whereas others are expressed at very low levels. The NDS profile (black) on a logarithmic scale makes it clear that almost every region in the genome is transcribed at some level. As anticipated, the DDS profile (green) for rightward and leftward transcripts combined on a logarithmic scale is similar to that for the NDS (black). The plots in the upper windows show the DDS profiles for rightward (magenta) and leftward (cyan) transcripts separately, on a logarithmic scale, and demonstrate that ASTs are produced from many regions. Further features of Fig. 1 are expanded on below.

Previously Unrecognized Transcripts. Selected transcribed regions that apparently lack CRs were investigated by using RACE (Fig. S1 and Table S1) and Northern blotting (NB) (Fig. S2). These experiments led to the annotation of one previously unrecognized CR (*US33A*), further characterization of a recently proposed CR (*UL30A*) (4), mapping of an NNT (RNA4.9), and investigation of

Author contributions: D.G., S.S., D.J.D., P.H., G.W.G.W., and A.J.D. designed research; D.G., S.S., C.C., M.H., D.J.D., K.B., R.D.H., J.G., and A.J.D. performed research; D.G., S.S., and G.W.G.W. contributed new reagents/analytic tools; D.G., S.S., R.D.H., P.H., G.W.G.W., and A.J.D. analyzed data; and D.G., G.W.G.W., and A.J.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequence reported in this paper has been deposited in the European Nucleotide Archive (www.ebi.ac.uk/ena/; accession no. ERA031126).

[1]Present address: Department of Chemistry, Biochemistry and Laboratory Medicine, Slovak Medical University, 833 03 Bratislava, Slovakia.

[2]Present address: Institute for Structural and Molecular Biology, Centre for Systems Biology at Edinburgh, University of Edinburgh, Edinburgh EH9 3JD, United Kingdom.

[3]To whom correspondence should be addressed. E-mail: andrew.davison@glasgow.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1115861108/-/DCSupplemental.

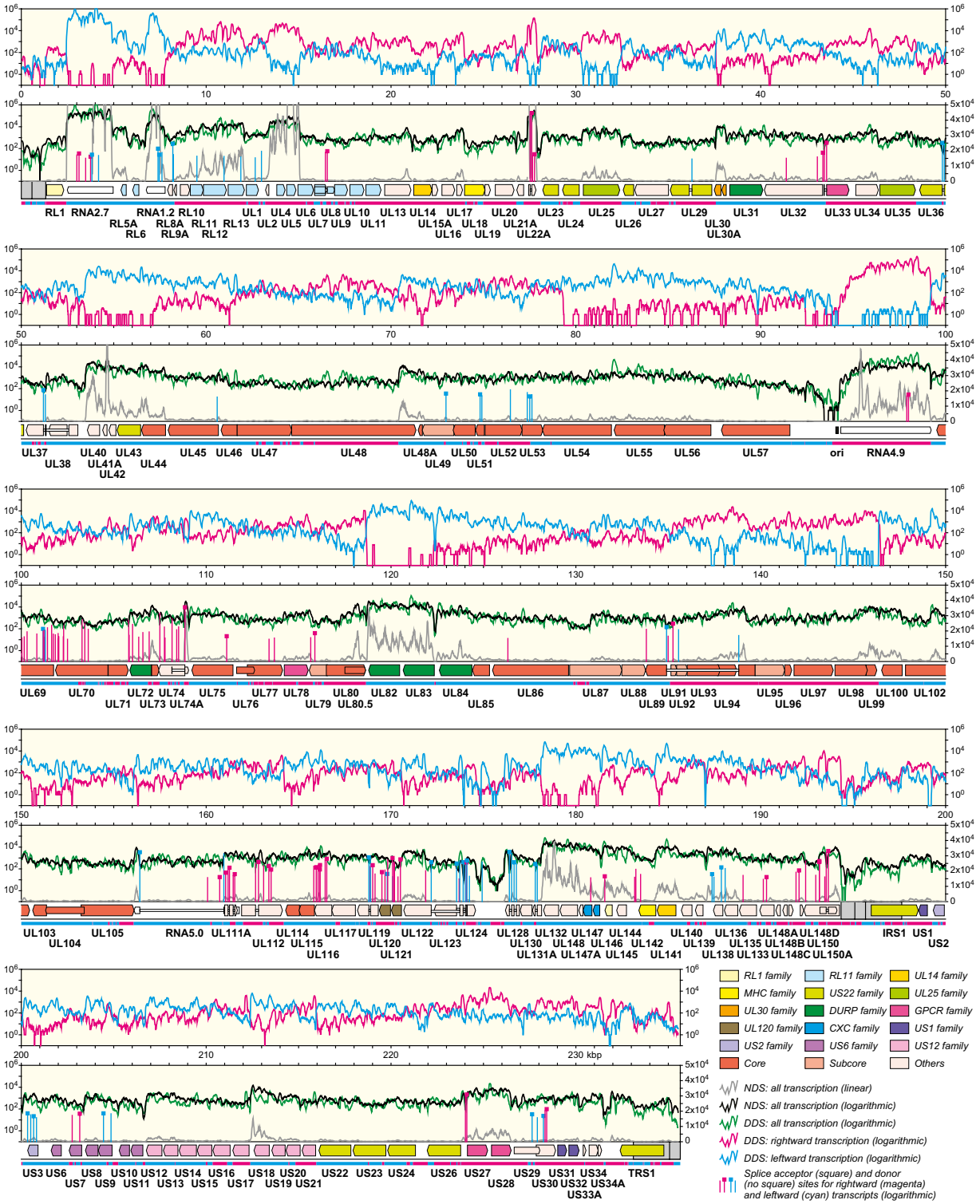


Fig. 1. Transcription and splicing of the HCMV genome. The genome is shown in five sections, with the inverted repeats (*ab*/*b′a′* and *a′c′*/*ca*) shaded gray. CRs are shown by color-shaded arrows and NNTs as white-shaded arrows, with gene nomenclature below. Introns connecting CRs or the *RNA5.0* exons are shown as narrow white bars. Colors applied to CRs indicate conservation among α-, β-, and γ-herpesviruses (core genes) or between β- and γ-herpesviruses (subcore genes), with subsets of the noncore genes grouped into families of related genes. *UL72* is both a member of the deoxyuridine triphosphatase-related protein (DURP) gene family and a core gene. The light yellow windows depict transcription profiles (*Methods*) as explained in the key in the bottom right corner. The plots in the lower windows show rightward and leftward transcripts combined: the NDS profile on linear and logarithmic scales (right and left, respectively) and the DDS profile on a logarithmic scale. The plots in the upper windows show the DDS profiles for rightward and leftward transcripts separately, on a logarithmic scale (right and left, respectively). The horizontal bar above the gene nomenclature shows regions where rightward (magenta) or leftward (cyan) transcription predominates. The vertical lines in the lower windows show the locations of the subset of splice sites supported by >10 reads, oriented rightward or leftward as explained in the key. The height of the line depicts the number of reads representing the site, plotted on a logarithmic scale (left). The complete list of splice sites is provided in Dataset S1.

ASTs. *US33A*, *UL30A* and *RNA4.9* are included in Fig. 1 and Fig. S3.

The size of the *US33A* transcript (1.3 kb) (Fig. S2) is consistent with the locations of the 5′ and 3′ ends (Table S1 and Fig. S3). A *US34A* probe also detected this RNA, plus major 0.9- and minor 0.4-kb transcripts (Fig. S2), suggesting that *US34* and *US34A* are 3′-coterminal with *US33A*. *US33A* is conserved in chimpanzee cytomegalovirus (CCMV), which is the closest relative of HCMV (19) (Fig. S4).

UL30A is related to the adjacent gene *UL30*, conserved among primate cytomegaloviruses, and potentially expressed from a nonconventional initiation codon (ACG) (4). The location of the 3′ end in the middle of *UL30* (Table S1 and Fig. S3) is consistent with the major 0.55-kb transcript (Fig. S2). Minor RNAs of 5.7 and 3.5 kb were also apparent, the latter perhaps representing a *UL32* transcript that is 3′-coterminal with *UL30A* and antisense to *UL31*.

The 5′ and 3′ ends of RNA4.9 (Table S1 and Fig. S3) accord with the transcript size of 4.9 kb (Fig. S2). The location of the 3′ end of a major RNA apparently mapped from partial transcripts is the same as that of this NNT (15, 20, 21). The majority (65.1%) of viral polyA RNA transcription is committed to production of NNTs, with 46.8, 7.9, 10.1, and 0.3% of appropriately oriented DDS reads representing RNA2.7, RNA1.2, RNA4.9, and RNA5.0, respectively. The RNA2.7, RNA1.2, RNA4.9, and RNA5.0 populations comprise 65.7, 26.9, 7.2, and 0.2% of NNT molecules, respectively. These statistics prompt the startling observation that only about one-third of viral polyA RNA transcription is protein-coding. The most abundant mRNA is that of *UL22A*, which encodes an immunoregulatory RANTES (regulated upon activation, normal T cell expressed and secreted) decoy receptor (22) and accounts for 1.7% of appropriately oriented DDS reads, equivalent to 14% of NNT molecules.

Extensive overlap of protein-coding RNAs specified by opposing strands of the HCMV genome is uncommon. Examples examined include the 3′ ends of *UL48A* and *UL53*, which are located well inside *UL48* and *UL54*, respectively (Table S1 and Fig. S3). In contrast, ASTs are apparent in most regions (Fig. 1). Overall, they account for 8.7% of transcription from CRs and are thus generally expressed at lower levels than their sense counterparts, although there are exceptions (e.g., *UL29*). RACE data indicate that one AST (UL115as) initiates at a specific location, and that several terminate downstream from polyA signals, with instances of a tail-to-tail arrangement in relation to other transcripts (Table S1). Examples (followed by the sense gene and the distance between polyA signals on the two strands) are UL16as (putative *UL15A*, 35 b), UL30as (*UL30A*, 36 b), UL121as (*UL122*, overlapping), UL145as (not arranged tail-to-tail), UL142as (*UL141*, 193 and 245 b), UL139as (*UL138*, 28 b), and IRS1as (*UL150A*, 1219 b).

Characterization of Splicing Patterns. The primary DS used for splicing analysis was the NDS, rather than the DDS, because fewer biochemical manipulations were involved in its generation. Custom software, capable of coping with genomes that have a high transcript density, was developed to identify reads containing potential splice junctions, from which donor and acceptor sites were proposed. These sites were required to be located adjacent to the canonical GT or AG dinucleotides, respectively, for the following reasons. Processing of RNA samples for deep sequencing involved reverse transcription, during which template switching may occur. This process involves transfer of an elongating cDNA molecule from one template RNA molecule to another, by repriming from a short sequence repeated in both templates (23). Template switching has been acknowledged to generate artifactual junctions in experimental studies (24). Also, the vast majority (~99%) of eukaryotic splice sites comply with the GT/AG rule (25), all previously recognized HCMV sites (described below) comply, and only a single noncompliant HCMV site was identified in ancillary searches (see below). Exerting the GT/AG rule resulted in the identification of 264 potential donor and 178 acceptor sites.

The number of NDS reads supporting every possible junction between these sites (regardless of order and orientation in the genome, a total of 46,992 combinations) was determined. A heuristic approach was taken to exclude junctions that most probably arose from template switching and happened to comply with the GT/AG rule. Excluded junctions were required to have at least two properties consistent with template switching: originating from highly transcribed regions (especially *RNA2.7*), supported by low read numbers, not involved in alternative splicing (a site that is alternatively spliced is much less likely to have been identified as a result of template switching than one that is not), characterized by repeats at the cognate genome locations, and not matching the wider consensuses for splice site sequences (25). This analysis led to the identification of 389 junctions supported by one or more reads, generated from 229 potential donor and 132 acceptor sites. Dataset S1 provides full information on these sites, and Fig. 1 shows the locations of the subset of sites supported by >10 reads. A total of 63 donor sites is located in CRs, 115 in ASTs, and 51 between CRs or in regions encoding NNTs. The corresponding numbers for the acceptor sites are 69, 31, and 32, respectively. Overall, 58 CRs contain one or more splice sites on the appropriate strand. The notation (e.g., D+27528) used below to denote a splice site is D (donor) or A (acceptor), + (rightward) or – (leftward), and the nucleotide location of the exon end. A solidus (/) indicates alternative sites, and a caret (^) indicates a junction.

To confirm splice junctions identified from the NDS analysis, a large number (96) was tested by RT-PCR and sequencing (Fig. S1). These were represented by 16–131,826 reads (Table S2 and Dataset S1), a range that extends over nearly four orders of magnitude. A total of 89 junctions was confirmed, as indicated by the colored borders of the cells in Dataset S1. For controls, two junctions classed as having arisen from template switching were tested and found negative (Table S2). RACE experiments carried out to map RNA ends (see below) incidentally confirmed many junctions, including two of the seven not detected by RT-PCR (Dataset S1). Moreover, three of the five not detected by RT-PCR or RACE involved alternatively spliced sites that were unlikely to have been identified as a result of template switching. This left 2 of the 96 tested junctions (D+3005^A+3122 and D+97900^A+97993) unsupported by RT-PCR and RACE and not involved in alternative splicing. Overall, only 26 pairs of sites identified by the NDS analysis were neither involved in alternative splicing nor tested by RT-PCR or RACE. From these considerations, it seems likely that the inferred splicing patterns are a close representation of reality.

Ancillary searches aimed at detecting sites that do not conform to the GT/AG rule but are alternatively spliced to conforming sites yielded a single example. This was D+192592 (as D+192592^A+193169/A+193172), in which the GT dinucleotide is substituted by GC, which is the most common noncanonical donor site (25).

Splice sites identified from the DDS accorded generally with those identified from the NDS, although there was variation in supporting read numbers (Dataset S1). Also, each DS yielded a unique set of sites represented by low read numbers (usually 1). These differences may have been due to the precise stage at which RNA was harvested, the conditions used for sample processing, and stochastic effects occurring during reverse transcription of rare RNAs.

Previously Reported Splicing Patterns. Studies on various HCMV strains have identified splice sites that are crucial for full-length protein expression. They are located in *UL22A* (10), *UL29* (5), *UL33* (26), *UL36* (27), *UL37* (27), *UL89* (10), *UL111A* (28), *UL112* (29), *UL119* (10), *UL122* (30), *UL123* (31, 32), *UL128* (33), and *UL131A* (33). All were confirmed in the present study. Moreover, many reported sites that are not crucial were also confirmed. Examples are those within *UL112* (29), *US3* (34), and *UL122* (30, 35), and those linking *UL122* and *UL119* (10). The

sites in *RNA5.0* (10) were also identified. However, other reported sites were not detected. Examples are those within *RNA2.7* (10), *UL37* (36), and *UL123* (IE19 and IE17.5) (37, 38), and most instances in the *UL115–UL119* region (39). Lack of detection of these sites does not imply that the original identifications were erroneous, as they may have been expressed at very low levels or at other stages of the replication cycle, or they may not be conserved in strain Merlin.

Previously Unrecognized Splicing Patterns. The NDS analysis detected splice sites sensitively, but was not capable of determining the extent of exons or the ends of RNAs. These features were investigated in selected transcripts by using RACE (Fig. S1 and Table S1) and NB (Fig. S2), which are less sensitive and identified major species only. Key findings are summarized in Fig. 1 and Fig. S3.

Splicing in two genes (*UL8* and *US27*) was revealed, using sites that are conserved in CCMV. These sites had been predicted previously, but experimental confirmation was unsuccessful (10). *UL7* and *UL8* are members of the RL11 gene family (Fig. 1), and encode predicted membrane glycoproteins (40). An intron (D+16449^A+16552) links *UL7* and *UL8*, thus 5′-extending the original *UL8*, with *UL7* being expressed from the unspliced RNA (Fig. 1 and Dataset S1). The expression ratio (ER) of spliced to total RNA for *UL8* was 0.57, the remainder presumably constituting the unspliced *UL7* transcript. *US27* is a member of the GPCR (G protein-coupled receptor) gene family (Fig. 1). An intron (D+224024^A+224100) is apparent in the 5′ leader (ER = 0.85) (Fig. S3 and Dataset S1). RACE and NB experiments confirmed that the *US27* and *US28* transcripts (2.7- and 1.3-kb, respectively) are 3′-coterminal (Table S1 and Fig. S2).

Five acceptor sites are notable in being spliced from one or another of a large number of upstream exons: A–8196 (25 exons), A+108854 (46 exons), A+135253 (22 exons), A+174081 (9 exons), and a pair at A+193169/A+193172 (11 exons) (Dataset S1). Excepting one of the pair, these superacceptor sites (SASs) are conserved in CCMV, and each is located upstream from a potential CR (*RL8A*, *UL74A*, *UL92*, *UL124*, and *UL150A*, respectively) (Fig. S3). The observation that *RL8A*, *UL92*, *UL124*, and *UL150A* contain a potential ATG initiation codon suggests that upstream exons may form untranslated leaders, although this does not rule out the possibility that some may 5′-extend the CR. In contrast, *UL74A* lacks an initiation codon, and must be supplied with one by at least one upstream exon. Because SASs form an unusual class of acceptor site, four of the five examples were investigated in further detail.

The partial transcript arrangement in the vicinity of the SAS at A–8196 has been deduced previously (12, 40, 41), and was extended to include transcripts from previously unrecognized genes *RL8A*, which is downstream from the SAS, and *RL9A* (Fig. 1 and Fig. S3). Both predicted proteins contain potential transmembrane regions and are conserved in CCMV (Fig. S4). In RACE experiments, four 5′ ends of upstream exons spliced to *RL8A* were identified, and two 5′ ends for *RL9A*, which is not spliced, were mapped (Table S1 and Figs. S1 and S3). Both genes share a 3′ end with *RNA1.2* at 6755. They appear to be expressed at low levels, as transcripts of the sizes inferred from RACE were not detected by NB. The 4.2-kb transcript detected using an *RL8A* probe (Fig. S2) is too large and of unknown provenance, and no *RL9A* transcripts were detected.

The *UL74A* CR is preceded by the SAS at A+108854, conserved among β-herpesviruses, and encodes a predicted membrane glycoprotein (3). Four 5′ ends of upstream exons were identified by RACE, and the 3′ end mapped to 109088 (Table S1 and Figs. S1 and S3). These locations are consistent with the major 0.8-kb transcript (*UL73*; ER = 0.96) and heterogeneous 0.45-kb transcripts (*UL74A*; ER = 0.28) detected by NB (Fig. S2). The origin of the minor 1.5-kb RNA is unknown. Some of the features of *UL74A* in strain Merlin are equivalent to those for the corresponding exon (*UL73.5*) in strain AD169 (3). In particular, *UL74A* acquires an initiation codon from the upstream exon spliced at D+108155 (Fig. 1 and Fig. S3). It is not known whether initiation codons are also supplied by other upstream exons, thus generating *UL74A* proteins with heterogeneous N termini. Many of the upstream exons, and others downstream from *UL74A*, are alternatively spliced to the SAS at A+135253, which is upstream from a recognized ORF, *UL92* (Dataset S1). This SAS was not characterized further.

A single 5′ end of an upstream exon spliced to the SAS at A+174081, which precedes *UL124*, was identified, and the 3′ end mapped to 174649 (Table S1 and Figs. S1 and S3). *UL124* is conserved in CCMV and encodes a predicted membrane glycoprotein (19). The major 0.7-kb transcript detected by NB is consistent with these locations (Fig. S2). A minor 5.5-kb transcript of unknown provenance was also apparent.

Four 5′ ends of upstream exons spliced to the SASs at A+193169/A+193172 were identified by RACE, and the 3′ end mapped to 194348 (Table S1 and Figs. S1 and S3). The two protein-coding exons of previously unrecognized gene *UL150A* (ER = 0.97), separated by an intron (D+193542^A+193654), are located downstream from the SASs (Fig. 1 and Fig. S3). The splice sites between the protein-coding exons are conserved in CCMV, as is the encoded protein (Fig. S4). The intervening intron specifies one of the HCMV miRNAs (miR-UL148D) (17). Sizes at the lower end of the heterogeneous transcripts (1.2–1.6 kb) detected by NB are consistent with the RACE data (Fig. S2). The *UL150A* CR is unusual in that it is entirely antisense to a larger, unspliced CR (*UL150*; 7.5-kb RNA) (Fig. 1 and Figs. S2 and S3), which is also conserved in CCMV (19). In the region of overlap, the amino acid sequence of *UL150A* is better conserved (47% identical) than that of *UL150* (40%), and codon bias is marginally more favorable (0.94 and 0.93, respectively, using the *UL54*, *UL82*, and *UL86* CRs for calibration). The prediction of *UL150A* as a CR was tested by expressing V5-tagged *UL150A* proteins in the context of HCMV infection. A doublet at ∼34 kDa and a minor species at 50 kDa were detected at 1, 3, and 6 d after infection (Fig. 2). The predicted mass of the tagged primary translation product is 31 kDa. The detection of more than one protein might reflect protein processing or splicing from upstream exons that 5′-extend the CR.

The NDS analysis also yielded evidence for long-range splicing, some of which was supported by RACE data, for example, between *US9* and the *RL8A* SAS (D–204854^A–8196) (Dataset S1). Moreover, splicing was detected between exons that are located out of order in, or on different strands of, the genome. For example, some of the exons spliced to the *RL8A* or *UL150A* SASs were alternatively spliced to the *UL74A* SAS (Dataset S1). These junctions may have resulted from long-range transcription of concatemeric genomes, in which U_L and U_S are present in either orientation.

Discussion

A number of technical limitations apply to the deep-sequencing approach used. Certain spliced polyA RNAs would not have been detected: those containing very short exons, and those not

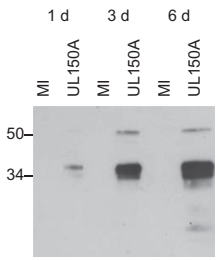


Fig. 2. Immunoblot analysis of V5-tagged *UL150A* proteins expressed in the context of HCMV. Proteins harvested at 1, 3, and 6 d after infection from mock-infected cells (MI) or cells infected with the HCMV recombinant (UL150A) were probed using a V5-Tag antibody. Estimated masses (kDa) are shown.

originating via alternative splicing and either consisting of exons mapping <50 b or >32 kb apart in the genome or using noncanonical splicing. Moreover, the data did not reliably facilitate the characterization of intact transcripts, including the locations of 5′ and 3′ ends. Points of interest in this regard were tested by RT-PCR, RACE, and NB. Also, nonpolyA transcripts (including miRNAs) would not have been represented.

The information arising from the study is as follows. Transcripts encoded by four previously unrecognized genes (*RL8A*, *RL9A*, *UL150A*, and *US33A*) were mapped, and the *UL150A* CR, which completely overlaps that of *UL150*, was shown to be expressed as protein during virus infection. The number of NNTs was expanded to four (RNA2.7, RNA1.2, RNA4.9, and RNA5.0). ASTs were detected throughout the genome, some having discrete 5′ and 3′ ends and some mapping tail-to-tail with other RNAs. Splicing is more common than recognized previously, and there is evidence for long-range transcription. *UL8* and *UL150A* were added to the genes comprising multiple coding exons. Alternative splicing is frequent, with four genes (*RL8A*, *UL74A*, *UL124*, and *UL150A*) and perhaps one other (*UL92*) expressed via SASs.

These findings demonstrate that HCMV transcription is more complex and multifaceted than recognized previously, with more CRs, splicing, NNTs, and ASTs having become apparent. This has implications for the potential sophistication of HCMV functionality during infection. Approximately one-third (58/170) of HCMV CRs contain splice sites, and the extent to which this affects viral protein-coding capacity is now a question of great interest. Some spliced RNAs are clearly crucial for expression of full-length proteins, and others that are not so obviously in this category may nonetheless have important roles. However, other spliced RNAs may be nonfunctional, representing transcripts that undergo a degree of splicing merely because their synthesis happens to traverse splice sites. This may apply to spliced RNAs that are very rare in comparison with their unspliced counterparts (e.g., ER = 0.003 for D–181199^A–181096 in *UL146*). Members of one class of noncoding RNAs, NNTs, are encoded by dedicated regions of the genome and likely to have important roles, and more examples probably await discovery. A second class of noncoding RNAs, ASTs, has its historical origins in antiparallel (symmetric) transcripts, which were first reported several decades ago in herpes simplex virus (42). The most extensive study of HCMV to date (15) used cDNA cloning of strain AD169 transcripts to detect ASTs in 34 of the CRs annotated in strain Merlin, and three (UL24as, UL36as, and UL102as) were characterized further by NB. The DDS analysis indicates that ASTs are associated with most CRs (Fig. 1), and that complex splicing is evident in some (e.g., the *UL114-UL121* region, which contains UL115as) (Table S1 and Figs. S1 and S3). The functions of ASTs are unknown, and it is possible that some ASTs result from transcriptional profligacy and have none. However, the observation that the 3′ ends of some ASTs appear to have been under evolutionary constraint in relation to neighboring genes implies that they may be important. If so, roles can only be guessed at this stage, with one obvious possibility being to regulate sense strand expression. Some ASTs may exert their effects via cleavage into miRNAs (which represent another layer of genomic complexity) or translation into short oligopeptides (43). However, in these models only a very minor proportion of AST content would be functional, leaving the role of the major proportion in question. One practical outcome of HCMV transcript complexity is that studies using mutants should be conducted with a heightened degree of circumspection.

Deep sequencing represents a stage in HCMV genomics that complements and extends previous approaches, such as microarray-based analyses (44). We have demonstrated the capability of this approach to provide high-resolution information on HCMV transcription and splicing. It can be extended readily to studies of HCMV transcription in other strains and cell types, expression kinetics during lytic and latent infections, comparative properties of protein-coding and noncoding RNAs, and transcription in vivo. It can also be applied to the genomics of other nuclear DNA viruses.

Methods

Preparation of RNA. HCMV strain Merlin, which is stable in vitro by virtue of point mutations in two genes (*RL13* and *UL128*) (45), was grown at 37 °C in human fetal foreskin fibroblast cells (HFFFs; HFFF2, European Collection of Cell Cultures 86031405) in DMEM containing 10% FCS. HFFF2 monolayers (75-cm^2) were infected with 5 pfu per cell of HCMV and incubated for 72 h. Total infected cell RNA was isolated using TRI reagent (Sigma) or an RNeasy Midi Kit (Qiagen). Three preparations were made, one in 2005 and two in 2010.

Deep Sequencing. Data were generated by using an Illumina Genome Analyzer IIx in The Sir Henry Wellcome Functional Genomics Facility (University of Glasgow). For the NDS, an Illumina mRNA-Seq prep kit was used: polyA RNA was selected from 1 μg of the 2005 RNA preparation, fragmented, and subjected to first- and second-strand cDNA synthesis by using random primers. The DNA was then end-repaired, 3′-adenylated, ligated to adapters, PCR-amplified, and size-selected (mode = 245 bp). For the DDS, Illumina mRNA-Seq prep and small RNA sample prep kits were used: polyA RNA was selected from 2 μg of one of the 2010 RNA preparations, fragmented, phosphatased, phosphorylated, ligated to RNA adapters, subjected to first-strand cDNA synthesis and PCR-amplified by using adapter-specific primers, and size-selected (mode = 188 bp).

Transcription Levels. Unaligned, 76-b reads with associated phred quality scores were generated by using SCS v2.6 or v2.8, RTA v1.12, and CASAVA v1.7 (Illumina), and stored in fastq files. The NDS and DDS were then assembled against a reference sequence by using Maq v0.7.1 under default alignment settings (46). The assembly was visualized by using Tablet v1.10.05.21 (47). The reference consisted of the strain Merlin genome sequence (GenBank accession no. AY446894.2) (2) or a truncated version lacking all but 50 b of the terminal repeats (*ab* and *ca*). DDS reads were also separated into DSs representing rightward and leftward transcripts and assembled independently. Transcription profiles were calculated from the number of reads commencing in 10-b windows and summing these over 100-b windows (10-b increment for each).

Splice Site Detection. A primary DS (NDS or DDS) was processed to gather reads containing potential splice junctions by the following steps. The rationale was that at least 21 b of the 5′ and 3′ parts of a junction-containing read must match two regions at least 50 b apart in the HCMV genome. Perl scripts are available from the authors.

i) Duplicate reads were removed, and a regular expression (regexp) search was conducted for reads that contain exact matches between the 21 b at the 5′ and 3′ ends (positions 1–21 and 56–76) and two different regions in the reference, and that meet the following criteria: (*i*) matches on the same strand of the reference, (*ii*) order of matching sequences in the read preserved in the reference, and (*iii*) matching sequences separated in the reference by 50 b to 32 kb.

ii) Each read was divided into two after position 21, and a regexp search was conducted for exact matches of the entire 5′ and 3′ parts to the reference. If criteria *i–iii* were met and the intervening genome sequence began with GT and ended with AG, a pair of 100 b exon–end sequences was generated. This procedure was repeated iteratively until position 55 was reached, to ensure that sequence redundancy at exon–ends did not result in failure to identify precise locations. Two sets of exon–end sequences were generated from the reads, one corresponding to donor sites and the other to acceptor sites.

iii) Step *ii* was repeated for the reverse complement of the reference, and the sets of exon–end sequences were combined with those sets from step *ii*.

iv) To avoid problems associated with lower-quality data toward the 3′ ends of reads, they were trimmed to 60 and 50 b and steps *i–iii* were repeated with appropriate modifications. The sets of exon–end sequences were combined with those from step *iii*, and duplicates were removed.

v) The exon–end sequences from step *iv* were joined to give a set of 200-b junctions comprising every donor site joined to every acceptor site in pairwise combinations. The central 20 b was extracted from each, and a sub-DS of reads specific to that sequence was generated by screening the primary DS for exact matches. Each DS was then assembled against the cognate 200-b sequence by using Maq. Thus, reads supporting each

junction were required to match the reference exactly in the 20-b sequence (i.e., 10 b from each exon) and substantially in the flanking regions (<3 b mismatch in 56 b).

ERs for selected spliced RNAs were calculated as described in step v, by extracting 20-b sequences (10 b each side of the spliced exon–exon junction or unspliced exon–intron junctions) from the DDS, generating read sub-DSs, assembling them against the cognate spliced or unspliced 200-b sequence, and counting the numbers of appropriately oriented reads. Complete splicing would yield an ER of 1.

RNA Mapping. RT-PCR, RACE, and NB were conducted using the 2010 RNA preparations. Titanium One-Step RT-PCR and SMARTer RACE Kits (Clontech) were used with primers listed in Tables S2 and S3, and products were sequenced directly or from plasmid clones. For NB, RNA was transferred from 1% formaldehyde-agarose gels to membranes, and hybridized with biotin-labeled, single-stranded RNA probes produced from plasmids (checked by sequencing) using a DIG Northern Starter Kit (Roche) and primers listed in Table S4.

Tagged Protein Expression. Published methods (45) were used to introduce a V5 epitope tag (checked by sequencing) at the C terminus of the *UL150A* protein by modifying bacterial artificial chromosome pAL1111, which contains the strain Merlin genome, and generating a recombinant virus (Table S5). HFFFs were mock-infected or infected at 5 pfu per cell and incubated in FCS-free DMEM for 1 or 2 d, or in DMEM containing 10% FCS for 3 d and then FCS-free DMEM for 3 d. Cell extracts were electrophoresed in 10% SDS/polyacrylamide gels, transferred to membranes, probed with a V5-Tag mouse monoclonal antibody (MCA1360; AbD Serotec), and reprobed with horseradish peroxidase-conjugated goat anti-mouse IgG (170-6516; Bio-Rad). Signal was detected by using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

ACKNOWLEDGMENTS. We thank Wai Kwong Lee and Stewart Laing (British Heart Foundation Glasgow Cardiovascular Research Centre, University of Glasgow) for DNA sequencing services and Elaine Douglas and Fiona Jamieson for technical assistance. This work was funded by the UK Medical Research Council and The Wellcome Trust.

1. Davison AJ (2007) *Human Herpesviruses: Biology, Therapy and Immunoprophylaxis*, eds Arvin A, et al. (Cambridge University Press, Cambridge, UK), pp 10–26.
2. Dolan A, et al. (2004) Genetic content of wild-type human cytomegalovirus. *J Gen Virol* 85:1301–1312.
3. Scalzo AA, Forbes CA, Smith LM, Loh LC (2009) Transcriptional analysis of human cytomegalovirus and rat cytomegalovirus homologues of the M73/M73.5 spliced gene family. *Arch Virol* 154(1):65–75.
4. Davison AJ (2010) Herpesvirus systematics. *Vet Microbiol* 143(1):52–69.
5. Mitchell DP, Savaryn JP, Moorman NJ, Shenk T, Terhune SS (2009) Human cytomegalovirus UL28 and UL29 open reading frames encode a spliced mRNA and stimulate accumulation of immediate-early RNAs. *J Virol* 83:10187–10197.
6. Bego M, Maciejewski J, Khaiboullina S, Pari G, St Jeor S (2005) Characterization of an antisense transcript spanning the UL81-82 locus of human cytomegalovirus. *J Virol* 79: 11022–11034.
7. Murphy E, Shenk T (2008) Human cytomegalovirus genome. *Curr Top Microbiol Immunol* 325:1–19.
8. Varnum SM, et al. (2004) Identification of proteins in human cytomegalovirus (HCMV) particles: The HCMV proteome. *J Virol* 78:10960–10966.
9. McDonough SH, Staprans SI, Spector DH (1985) Analysis of the major transcripts encoded by the long repeat of human cytomegalovirus strain AD169. *J Virol* 53:711–718.
10. Rawlinson WD, Barrell BG (1993) Spliced transcripts of human cytomegalovirus. *J Virol* 67:5502–5513.
11. Kulesza CA, Shenk T (2004) Human cytomegalovirus 5-kilobase immediate-early RNA is a stable intron. *J Virol* 78:13182–13189.
12. Hutchinson NI, Tocci MJ (1986) Characterization of a major early gene from the human cytomegalovirus long inverted repeat; predicted amino acid sequence of a 30-kDa protein encoded by the 1.2-kb mRNA. *Virology* 155(1):172–182.
13. Reeves MB, Davies AA, McSharry BP, Wilkinson GWG, Sinclair JH (2007) Complex I binding by a virally encoded RNA regulates mitochondria-induced cell death. *Science* 316:1345–1348.
14. Kulesza CA, Shenk T (2006) Murine cytomegalovirus encodes a stable intron that facilitates persistent replication in the mouse. *Proc Natl Acad Sci USA* 103:18302–18307.
15. Zhang G, et al. (2007) Antisense transcription in the human cytomegalovirus transcriptome. *J Virol* 81:11267–11281.
16. Ma YP, et al. (2011) Novel transcripts of human cytomegalovirus clinical strain found by cDNA library screening. *Genet Mol Res* 10:566–575.
17. Pfeffer S, et al. (2005) Identification of microRNAs of the herpesvirus family. *Nat Methods* 2:269–276.
18. Prichard MN, et al. (1998) Identification of persistent RNA-DNA hybrid structures within the origin of replication of human cytomegalovirus. *J Virol* 72:6997–7004.
19. Davison AJ, et al. (2003) The human cytomegalovirus genome revisited: Comparison with the chimpanzee cytomegalovirus genome. *J Gen Virol* 84(Pt 1):17–28.
20. Scott GM, Barrell BG, Oram J, Rawlinson WD (2002) Characterisation of transcripts from the human cytomegalovirus genes TRL7, UL20a, UL36, UL65, UL94, US3 and US34. *Virus Genes* 24(1):39–48.
21. Davis MG, Huang ES (1985) Nucleotide sequence of a human cytomegalovirus DNA fragment encoding a 67-kilodalton phosphorylated viral protein. *J Virol* 56(1):7–11.
22. Wang D, Bresnahan W, Shenk T (2004) Human cytomegalovirus encodes a highly specific RANTES decoy receptor. *Proc Natl Acad Sci USA* 101:16642–16647.
23. Temin HM (1993) Retrovirus variation and reverse transcription: Abnormal strand transfers result in retrovirus genetic variation. *Proc Natl Acad Sci USA* 90:6900–6903.
24. Houseley J, Tollervey D (2010) Apparent non-canonical trans-splicing is generated by reverse transcriptase in vitro. *PLoS One* 5:e12271.
25. Sheth N, et al. (2006) Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res* 34:3955–3967.
26. Margulies BJ, Browne H, Gibson W (1996) Identification of the human cytomegalovirus G protein-coupled receptor homologue encoded by UL33 in infected cells and enveloped virus particles. *Virology* 225(1):111–125.
27. Kouzarides T, Bankier AT, Satchwell SC, Preddy E, Barrell BG (1988) An immediate early gene of human cytomegalovirus encodes a potential membrane glycoprotein. *Virology* 165(1):151–164.
28. Kotenko SV, Saccani S, Izotova LS, Mirochnitchenko OV, Pestka S (2000) Human cytomegalovirus harbors its own unique IL-10 homolog (cmvIL-10). *Proc Natl Acad Sci USA* 97:1695–1700.
29. Wright DA, Spector DH (1989) Posttranscriptional regulation of a class of human cytomegalovirus phosphoproteins encoded by an early transcription unit. *J Virol* 63: 3117–3127.
30. Stenberg RM, Witte PR, Stinski MF (1985) Multiple spliced and unspliced transcripts from human cytomegalovirus immediate-early region 2 and evidence for a common initiation site within immediate-early region 1. *J Virol* 56:665–675.
31. Stenberg RM, Thomsen DR, Stinski MF (1984) Structural analysis of the major immediate early gene of human cytomegalovirus. *J Virol* 49(1):190–199.
32. Akrigg A, Wilkinson GWG, Oram JD (1985) The structure of the major immediate early gene of human cytomegalovirus strain AD169. *Virus Res* 2(2):107–121.
33. Akter P, et al. (2003) Two novel spliced genes in human cytomegalovirus. *J Gen Virol* 84:1117–1122.
34. Weston K (1988) An enhancer element in the short unique region of human cytomegalovirus regulates the production of a group of abundant immediate early transcripts. *Virology* 162:406–416.
35. Kerry JA, et al. (1995) Isolation and characterization of a low-abundance splice variant from the human cytomegalovirus major immediate-early gene region. *J Virol* 69: 3868–3872.
36. Adair R, Liebisch GW, Colberg-Poley AM (2003) Complex alternative processing of human cytomegalovirus UL37 pre-mRNA. *J Gen Virol* 84:3353–3358.
37. Shirakata M, et al. (2002) Novel immediate-early protein IE19 of human cytomegalovirus activates the origin recognition complex I promoter in a cooperative manner with IE72. *J Virol* 76:3158–3167.
38. Awasthi S, Isler JA, Alwine JC (2004) Analysis of splice variants of the immediate-early 1 region of human cytomegalovirus. *J Virol* 78:8191–8200.
39. Leatham MP, Witte PR, Stinski MF (1991) Alternate promoter selection within a human cytomegalovirus immediate-early and early transcription unit (UL119–115) defines true late transcripts containing open reading frames for putative viral glycoproteins. *J Virol* 65:6144–6153.
40. Davison AJ, et al. (2003) Homology between the human cytomegalovirus RL11 gene family and human adenovirus E3 genes. *J Gen Virol* 84:657–663.
41. Greenaway PJ, Wilkinson GWG (1987) Nucleotide sequence of the most abundantly transcribed early gene of human cytomegalovirus strain AD169. *Virus Res* 7(1):17–31.
42. Kozak M, Roizman B (1975) RNA synthesis in cells infected with herpes simplex virus. IX. Evidence for accumulation of abundant symmetric transcripts in nuclei. *J Virol* 15 (1):36–40.
43. Xu Y, Ganem D (2010) Making sense of antisense: Seemingly noncoding RNAs antisense to the master regulator of Kaposi's sarcoma-associated herpesvirus lytic replication do not regulate that transcript but serve as mRNAs encoding small peptides. *J Virol* 84:5465–5475.
44. Chambers J, et al. (1999) DNA microarrays of the complex human cytomegalovirus genome: Profiling kinetic class with drug sensitivity of viral gene expression. *J Virol* 73:5757–5766.
45. Stanton RJ, et al. (2010) Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication. *J Clin Invest* 120: 3191–3208.
46. Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 18:1851–1858.
47. Milne I, et al. (2010) Tablet—Next generation sequence assembly visualization. *Bioinformatics* 26:401–402.