Heinrich Pette Institute

Leibniz Institute for Experimental Virology

Role of the cytomegalovirus early 1 protein isoforms during viral replication

Dissertation

submitted to the

Department of Chemistry

Faculty of Mathematics, Informatics and Natural Sciences

University of Hamburg

In fulfillment of the requirements

for the degree of

Doctor of Natural Sciences (Dr. rer. nat.)

by

Tim Schommartz

born in Rostock

Hamburg 2016

Prof. Dr. Wolfram Brune (first reviewer)

Prof. Dr. Ulrich Hahn (second reviewer)

Day of oral defense: April 15, 2016

This study was conducted between March 2012 and February 2016 at the Heinrich Pette Institute Leibniz Institute for Experimental Virology under the supervision of Prof. Dr. Wolfram Brune and Dr. Markus Perbandt.

Zwei Dinge sind zu unserer Arbeit nötig:
Unermüdliche Ausdauer und die Bereitschaft etwas,
in das man viel Zeit und Arbeit gesteckt hat, wieder wegzuwerfen.
(Albert Einstein)

Contents

1 Abstract

Herpesviruses have large and complex DNA genomes. The largest among the herpesviruses, the cytomegaloviruses, encode over 170 genes. Although most herpesviral gene products are expressed from unspliced transcripts, a substantial number of viral transcripts are spliced. Some viral transcripts are subject to alternative splicing, which leads to the expression of several proteins from a single gene.

The four isoforms of the Early 1 (E1) proteins of both the murine cytomegalovirus (MCMV) and the human cytomegalovirus (HCMV) are encoded by such alternatively spliced gene transcripts. These transcripts are derived from the MCMV M112/113 gene and its homolog in HCMV, UL112/113. Surprisingly little is known about these essential genes and only very recently the different E1 isoforms have begun to be studied individually. Functional analysis of the individual E1 proteins is difficult as deletion and nonsense mutagenesis, both common methods used in the generation of viral gene knockout mutants, affect several or all E1 isoforms at the same time.

Here I demonstrate that individual E1 gene products can be inactivated selectively by mutagenesis of the splice donor or acceptor site and by intron deletion or substitution mutagenesis. The expression of each of the four MCMV E1 protein isoforms was inactivated individually, and the requirement of each isoform for viral replication was analyzed in fibroblasts, endothelial cells, and macrophages. It has been shown that the largest E1 isoform (p87), but not the other E1 isoforms (p33, p36, and p38), are essential for MCMV replication in cell culture. Moreover, the presence of one of the two medium-sized E1 isoforms (p36 or p38) and the presence of intron 1, but not its specific sequence, are required for viral replication. By applying similar mutagenesis strategies to UL112/113, it was possible to compare the importance of individual E1 isoforms in both viruses.

To further characterize the function(s) of the E1 proteins, potential cellular interaction partners for the essential large MCMV E1p87 isoform have been identified. To this end a SILAC-based approach using affinity purification and mass spectrometry was applied. Several proteins of the family of heterogeneous nuclear ribonucleoproteins (hnRNP) were identified as potential interaction partners and further investigated.

2 Zusammenfassung

Die Genome der Herpesviren gehören zu den größten und komplexesten Genomen aller humanpathogenen Viren. Cytomegaloviren, die größten Vertreter der Herpesviren, kodieren dabei über 170 Gene von denen in den meisten Fällen ungespleißte Transkripte erzeugt werden. Durch neuere Studien wurde allerdings deutlich, dass eine nicht zu vernachlässigende Anzahl der cytomegaloviralen Transkripte gespleißt wird. Unter diesen findet man auch einige Transkripte, welche durch alternatives Spleißen entstehen. Ein Vorgang durch den ein einzelnes Gen mehrere Protein Isoformen kodieren kann.

Die vier viralen Early 1 (E1) Isoformen werden, sowohl bei dem murinen Cytomegalovirus (MCMV) als bei dem humanen Cytomegalovirus (HCMV), von solch alternativ gespleißten viralen Transkripten kodiert. In MCMV werden die E1 Proteine durch die M112/113 Genregion kodiert, in HCMV durch den homologen UL112/113 Genlokus. Obwohl diese Genregionen essentiell für die virale Replikation sind, ist erstaunlich wenig über sie bekannt. Erst in jüngster Zeit wurde damit begonnen, die Funktion(en) individueller E1 Isoformen zu studieren. Da jedoch Standardmethoden der Virusmutagense, wie die Konstruktion viraler *knockout* oder *nonsense* Mutanten, in der Regel alle E1 Isoformen betreffen, gestaltet sich die Untersuchung einzelner E1 Isoformen als äußerst schwierig.

Die in dieser Arbeit vorgestellte Methode ermöglicht das selektive Inaktivieren einzelner E1 Isoformen durch den gezielten Einsatz von Mutation der Spleißakzeptor- und Spleißdonorsequenzen; sowie durch Deletion von Introns und Substitutionsmutagense. Durch Einsatz dieser Techniken wurde der Einfluss der Inaktivierung einzelner E1 Isoformen auf die virale Replikation in Fibroblasten, Endothelzellen und Makrophagen untersucht. Dabei wurde deutlich, dass die größte E1 Isoform (p87), jedoch nicht die anderen drei E1 Isoformen (p33, p36 und p38), essential für die Replikation von MCMV in Zellkultur sind. Interessanterweise ist allerdings mindestens eine der mittelgroßen E1 Isoformen (p36 oder p38), sowie das erste Intron der M112/113 Genregion, unabhängig von dessen spezifischer Sequenz, zwingend für die virale Replikation erforderlich. Durch Verwendung analoger Methoden für die UL112/113 Genregion konnten ähnliche Effekte auch für die Inaktivierung einzelner HCMV E1 Isoformen beschrieben werden.

Zusätzlich wurden potentielle zelluläre Interaktionspartner für die große MCMV E1p87 Isoform mit Hilfe einer SILAC basierten Kombination von Affinitätsaufreinigung und Massenspektrometrie identifiziert. Die unter diesen potentiellen Interaktionspartnern enthalten Proteine aus der Gruppe der *heterogeneous nuclear ribonucleoproteins* (hnRNP) wurden dabei weiterführend untersucht.

3 Introduction

3.1 Cytomegaloviruses

3.1.1 Clinical relevance of cytomegalovirus infections

It is thought that the *Herpesviridae* co-evolved with their respective hosts for millions of years [1, 2]. This has led to their excellent adaptation to those hosts. As a result of this adaptation they have developed numerous counter mechanisms to avoid detection and eradication by the host's immune system.

Cytomegaloviruses (CMV) belong to the herpesviruses, with human cytomegalovirus (HCMV) being the human pathogen. The virus was first isolated in the 1950s and named after the characteristic enlargement of infected cells [3, 4].

The virus is transmitted either via body fluids such as saliva, blood, urine, semen or vaginal fluid [5]. It can also be transmitted intrauterine from the mother to the unborn child or postnatally via breastfeeding [6, 7]. The global seroprevalence ranges between 40 % and 90 % depending on the developmental and hygienic status of the region in question (FIG 1) [8]. Numbers given for CMV seroprevalence in Germany differ depending on the study and section of the population investigated [9]. A recent study in children in Germany describes an age-adjusted seroprevalence of 27.4 % with seroprevalence increasing with age [10].

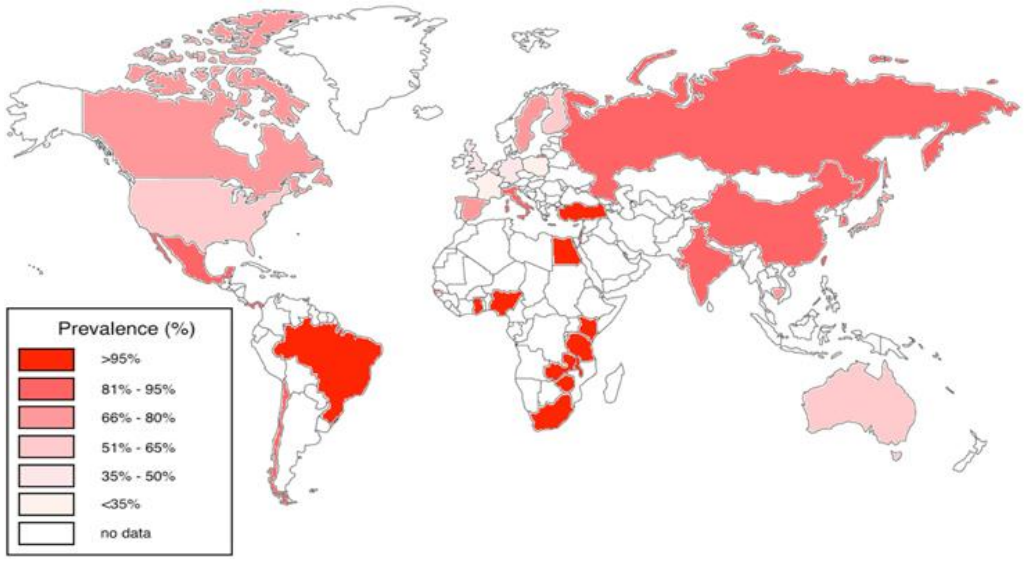


FIG 1. Worldwide CMV seroprevalence rates in adults. Modified from Adland et al., Ongoing burden of disease and mortality from HIV/CMV coinfection in Africa in the antiretroviral therapy era, Frontier in Microbiology, 2015

Infections are usually mild in immunocompetent individuals, sometimes accompanied by a CMV mononucleosis with mild flu-like symptoms [11]. Nevertheless, in rare cases more severe complications such as cardiovascular diseases have been reported upon primary infection with CMV [11-13].

While CMV infection is effectively controlled by the immune system in most healthy persons it can cause major complications in those that are immunocompromised. One such risk group is individuals infected with Human Immunodeficiency Virus (HIV) and suffering from Acquired Immune Deficiency Syndrome (AIDS). CMV infection in these patients may lead to severe complications such as retinitis and diseases of the gastrointestinal tract or central nervous system [14]. Fortunately, the outlook for those infected with HIV has improved. Whilst nearly every second AIDS patient was suffering from a CMV related complication in the past, these numbers dropped drastically with the availability of effective anti-retroviral therapies such as Highly Active Antiretroviral Therapy (HAART) [15]. However, organ transplant recipients, another high risk group, can develop complications similar to those experienced by untreated AIDS patients infected with CMV [16]. Those patients have the additional risk of the infection causing rejection of the transplanted organ [17]. The chance of this situation arising increases, if a seropositive organ is transplanted into a seronegative recipient, with the incidence of a CMV related complication ranging between 50 and 65 %. This chance is reduced if the recipient is already seropositive for CMV, although 15 to 20 % of all organ transplant recipients encounter CMV related complications [18].

Besides immunocompromised hosts CMV infections are also a major threat for immunologically immature hosts. This is especially the case in primary CMV infection of the mother during pregnancy, which can lead to the intrauterine infection of the unborn child. However, secondary infection of the mother can also lead to transmission of the virus to the child [19]. In industrialized countries congenital CMV infections are a main cause for long-term medical conditions in children. These conditions can be severe, ranging from sensory disorders - congenital CMV is the most common non-genetic cause of childhood hearing loss - inflammation of different organs, mental retardation or even death [19-22]. Astonishingly women's awareness of these severe threats is very low (FIG 2) [23-25]. There is also increasing evidence that CMV is a medical risk for the elderly, as well as immunocompetent persons in intensive care [26, 27].

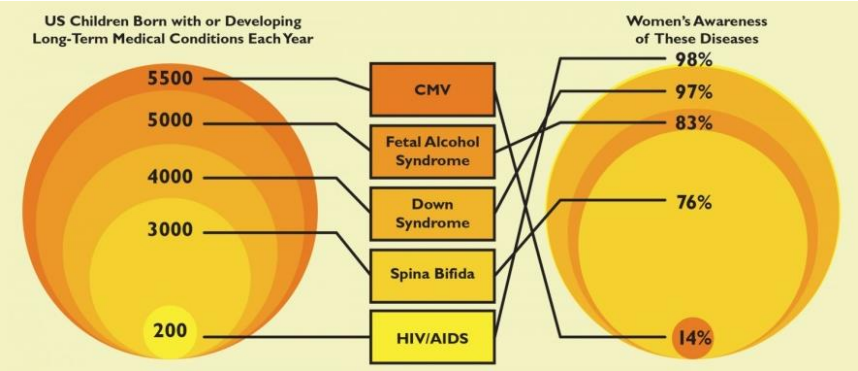


FIG 2. Number of children born with long term medical conditions in the US compared to the awareness of women in reproducible age of those conditions. Modified from http://photokapi.com

3.1.2 Classification and structure

Cytomegaloviruses are large DNA viruses and the most prominent members of the subfamily of β-*herpesvirinae* within the family *herpesviridae* from the order *herpesvirales* [28]. The α-, β- and γ-subfamily of *herpesviridae* are characterized by their host range, their site of latency and the timespan of their replicative cycle [29]. Cytomegaloviruses have a very narrow host range and can only infect cells of their natural host or very closely related species. Due to this limitation the murine cytomegalovirus (MCMV) is often used as a model system for HCMV infections in experimental settings [3, 30].

All CMVs have a biphasic lifecycle divided into lytic and latent phase. During the latent stage the virus persists lifelong in its host and no viral progeny is produced. This phase is sporadically interrupted by lytic phases where the full replicative cycle takes place and new viral particles are released [31]. The sites of latency are cells of the myeloid lineage, mainly monocytes and CD34 positive progenitor cells [32]. Conversely, the lytic phase is not as strictly restricted and can take place in a large number of cell types such as epithelial cells, endothelial cells, smooth muscle cells and fibroblasts [33, 34]. Compared to α-Herpesviruses, both HCMV and MCMV have a relatively long replicative cycle. While HCMV takes about 48 to 72 hours to complete a replicative cycle, MCMV replication only takes approximately 24 hours. At the end of the replicative cycle the host cell is lysed and the newly assembled viral particles are released [35].

CMV virions have a diameter ranging between 200 and 300 nm and carry the approximately 230 kb viral DNA genome enclosed in an icosahedral capsid [35, 36] (FIG 3). The capsid is enveloped by a plasma membrane studded with viral glycoproteins. The space in between the capsid and the membrane is termed the tegument. The

tegument is densely filled with viral and cellular proteins as well as cellular and viral RNA [35-38]. Both the HCMV as well as the MCMV genome have been predicted to code for 160 to more than 200 genes, with the number of possible gene products by far exceeding the number of genes [39-42].

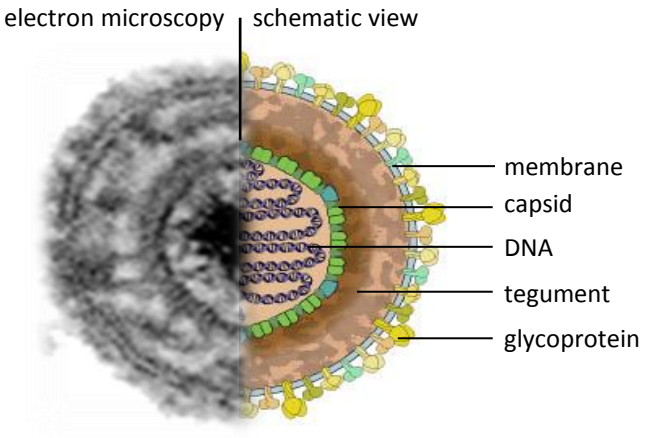


FIG 3. Structure of cytomegalovirus particles.
Merged view of electron microscopy picture and schematic drawing of a cytomegalovirus particle. Components of the virion are indicated in the figure. Modified from http://viralzone.expasy.org/all_by_species/180.html and http://www.virology.net/big_virology/bvdnaherpes.html website

3.1.3 Replicative cycle

The replicative cycle of CMV is a multi-step process (FIG 4) which begins with the attachment of virions to heparan sulfate proteoglycans at the host cell membrane [43]. This process is mediated by the viral glycoprotein gM and gB, with the latter being dispensable at least *in vitro* [44]. Attachment is followed by the binding of virus particles to a cellular receptor, a process mediated by the viral gH/gL protein complex [45-47]. The cellular receptor responsible for this is controversially discussed in the field with epidermal growth factor receptor, platelet-derived growth factor alpha receptor and several integrins being potential candidates [48-51].

After binding, viral particles enter the cell through a process often described as penetration [48, 50, 52]. The internalization of the virus differs between host cell types. In endothelial and epithelial cells, viral particles are internalized by endocytosis before being released into the cytoplasm by fusion of the viral envelope with the endosomal membrane [53]. In fibroblasts the viral envelope directly fuses with the cellular plasma membrane [54]. Upon fusion the viral capsids as well as the tegument components are

released into the cytoplasm, with the capsids then transported towards the nucleus along microtubules. There the viral DNA is imported into the nucleus where it circularizes and associates with the multipurpose cellular promyelotic leukemia protein-associated nuclear bodies (PML bodies), initiating viral gene expression [55-57].

Like all herpesviruses CMVs express their genes in a cascade-like fashion [35, 58] with the immediate early (IE) gene products being transcribed first. The IE genes mainly code for transcription factors and drive the expression of early (E) gene products [35]. The Early genes code for components of the viral replication machinery and their expression goes hand-in-hand with initiation of the "rolling circle" replication of viral DNA, starting at the origin of replication (termed *ori*Lyt) [59]. The E genes also trigger the expression of late (L) gene products, which are primarily structural proteins [60]. Late genes are further subdivided as either early late genes (E-L), which are expressed in low quantities before dramatically increasing in expression after the start of viral DNA replication, or as true late genes, which are exclusively expressed after the onset of viral DNA replication [35]. However, recent data based on mass spectrometry suggests subdividing the genes into five, instead of three, classes of viral gene products based on a finer temporal resolution of the expression of viral genes [61].

During late phases of infection the newly expressed capsid proteins are imported into the nucleus where they are assembled and packed with viral genomes. The fully assembled capsids then bud through the nuclear membranes and are transported along microtubules through the cytoplasm, with the tegument proteins accumulating around the capsid during transport [62, 63]. The tegument-surrounded capsids then receive their final envelope by budding into the Golgi apparatus associated secretory vesicles, which are highly enriched with viral glycoproteins, subsequently forming the outer viral lipid membrane. These vesicles are finally transported towards the plasma membrane of the cell where they fuse and release the newly formed virions [64, 65].

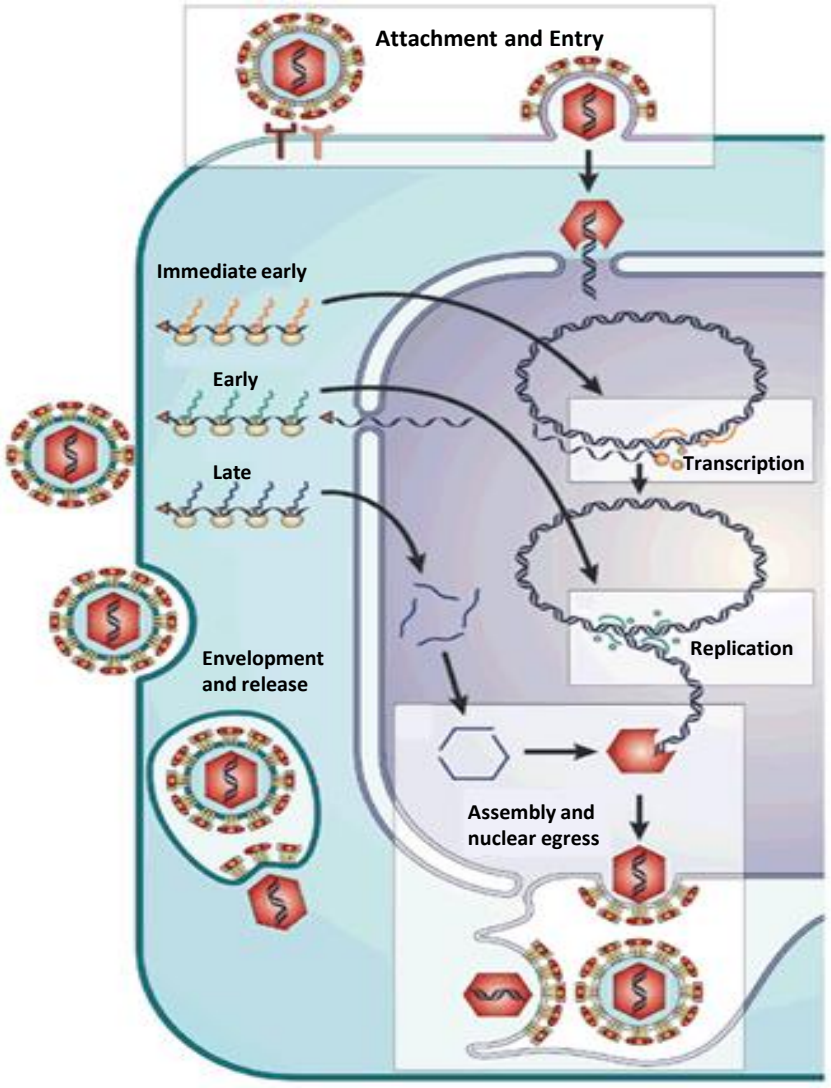


FIG 4. Cytomegalovirus replicative cycle. Binding of viral membrane proteins to cellular receptors on the plasma membrane of the cell leads to fusion of the two membranes and release of nucleocapsids containing the viral genome (red hexagons) into the cytoplasm. After being transported to nuclear pores the Viral DNA is liberated into the nucleus and circularizes. Cascade-like expression of three classes of viral gene products starts thereafter. Immediate-early proteins (orange) are involved in further transcription whereas early proteins (green) synthesize new viral DNA via “rolling circle” mechanism. Late proteins (blue) assemble into capsids, which incorporate newly replicated viral DNA. Nucleocapsids leave the nucleus by budding through the inner nuclear membrane into the perinuclear space. From there they undergo a secondary envelopment process before being released from the cell. Modified from: Coen and Schaffer, Antiherpesvirus drugs: a promising spectrum of new drugs and drug targets, Nature Reviews Drug Discovery, 2003.

3.1.4 The Cytomegalovirus E1 proteins

The CMV E1 proteins belong to the class of viral early proteins and exist in both HCMV and MCMV. In HCMV they are coded by the UL112/113 gene locus while in MCMV they are coded by M112/113, a positional homolog of UL112/113 [66, 67]. Both gene regions give rise to four protein isoforms derived from alternative splicing. These isoforms are phosphorylated and named after their apparent molecular mass [66-68]. For HCMV the E1 isoforms are E1p34, E1p43, E1p50 and E1p84 [67], whereas for MCMV the E1 isoforms are E1p33, E1p36, E1p38 and E1p87 [68]. The four E1 isoforms share a common N-terminus in both MCMV and HCMV [66-68].

For both viruses the E1 proteins are localized in the nucleus, where they are associated with the viral (pre-)replication compartments and form speckle-like structures[66, 69-71]. These speckle-like structures reduce over the course of the infection and become enlarged until they form big cloud-like structures in the nucleus [66, 69]. In the nucleus, both the HCMV E1 proteins and the MCMV E1 proteins are described to be associated with PML nuclear bodies [72, 73].

In general the CMV E1 proteins have major influence on viral replication. For HCMV, two studies report that inactivation of the UL112/113 gene products result in severely impaired viral replication in both strains Towne and AD169 [74, 75]. For MCMV it was shown that knockout of M112/113 results in a replication defective virus mutant [76]. Although this implies a crucial role of those proteins in CMV biology, surprisingly little is known about their function(s). Work published describing the role of UL112/113 and M112/113 is summarized below.

3.1.4.1 The HCMV UL112/113 gene products

The HCMV E1 proteins are thought to play a role in HCMV DNA replication. This was initially shown by co-transfection experiments using *ori*Lyt containing plasmids [77]. Together with ten other HCMV loci they serve as essential trans-acting factors for HCMV DNA replication. In a follow up study by Iskenderian and colleagues, it was shown that the UL112/113 gene products together with UL36-38, IRS1 or TRS1 and the major immediate early gene region activate expression of viral replication fork proteins which are essential for viral DNA replication [78]. The importance of the UL112/113 gene products for viral DNA replication was further strengthened by demonstrating that RNAi mediated reduction of the UL112/113 gene products significantly decreases viral DNA

replication [70]. Moreover it was shown that the HCMV E1 proteins colocalize with viral DNA [70].

There is evidence that this colocalization may be due to interaction of E1 proteins with (viral) DNA since it was shown in immunoprecipitation experiments the E1p43 and E1p50 bind to DNA-cellulose columns [69]. The E1p84 also seems to bind to the columns but to a much lower extent. Using chromatin immunoprecipitation it was shown that all four E1 proteins can bind to the *ori*Lyt in HCMV infected cells [79]. However, whether this is a direct interaction between the E1proteins and DNA or mediated by other cellular or viral factors has not been described yet.

Other studies present a more indirect link between the UL112/113 gene products and viral replication. It has been shown that the E1 proteins enhance the immediate early 2 (IE2) driven transactivation of the viral DNA polymerase (UL54) promotor [80, 81]. Transactivation of the UL54 promotor can also be mediated by E1p43 alone in the absence of IE2 [81]. This function of E1p43 is mediated by a specific domain spanning from amino acid 272 to 296. For enhancing the IE2 mediated transactivation of UL54 promotor an additional domain at amino acids 297 to 306 is needed [81]. Interestingly a later publication by Kim and colleagues showed interaction between all four E1 isoforms and IE2 if the isoforms were tested individually in transfection experiments [82]. This interaction appears to be mediated by the N-terminal part of the E1 proteins. However, whether or not the interaction of the isoforms E1p34, E1p50 and/or E1p84 with IE2 also leads to transactivation of the UL54 promotor has not been demonstrated yet.

The ability of UL112/113 gene products to function as transactivators is underlined by their ability to activate the full Kaposi's sarcoma-associated herpesvirus (KSHV) lytic cycle from latency [83].

Apparently the four E1 isoforms can also interact with each other and recruit the subunit of the HCMV DNA polymerase catalytic subunit (UL44) to viral pre-replication compartments [84]. Interaction among these proteins is important for their localization to the nucleus, especially for the translocation of E1p34 to the nucleus, and is mediated by the N-terminal part of the proteins [82, 84]. It was also shown that the E1 mediated translocation of UL44 depends on direct binding of E1p84 to UL44 [85]. This binding is required for the effective association with the *ori*Lyt. If the binding between UL44 and E1p84 is disrupted by deleting either the C-terminal part of E1p84 or the N-terminus and/or C-terminus of UL44, virus replication is severely impaired and *ori*Lyt dependent

DNA replication is significantly reduced [85]. Remarkably, disruption of the interaction among E1 proteins by deletion of parts of their N-terminus reduces their ability to recruit UL44 to nuclear pre-replication foci [82].

3.1.4.2 The MCMV M112/113 gene products

Most existing studies concentrate on the HCMV E1 proteins while only very few deal with the MCMV counterpart. One can assume that the MCMV E1 proteins share a number of properties with their homologs in HCMV, which are described in detail in paragraph 3.1.4.1. Nevertheless some publications also focus on the MCMV E1 proteins.

For those proteins it has been described that their expression is regulated by the immediate early 3 protein (IE3) which binds to a 10 bp activating motif (IE3AM) upstream of the M112/113 TATA box [86]. Upon binding of IE3 to IE3AM the M112/112 transcription is enhanced. The IE3AM is crucial for viral replication hence mutations in that sequence lead to drastically impaired MCMV replication [86]. Interestingly there is also an effect of the MCMV E1 proteins, most likely E1p87, on the expression of IE3 [87]. IE3 has been described to have a repressive effect on its own promotor the major immediate early promotor (MIEP) upon expression [88]. Apparently this auto-repression is relieved upon the expression of the M112/113 gene products [87].
The MCMV E1 proteins are also associated with viral persistence in the brain [89]. *In vivo* experiments by Arai and colleagues have shown that a reporter construct containing the M112/113 promotor is constitutively active in a subpopulation of neurons in the central nervous system (CNS) of mice. Upon infection the number of these neurons increased, providing evidence that this subpopulation of neurons may be the place of MCMV persistence in the brain [89]. Therefore the MCMV E1 proteins have also been used as markers for viral latency to study MCMV infection in the brain [90].

The overall importance of this gene region for MCMV has been additionally supported by the finding that mutations in the M112/113 locus enable the virus to overcome the species barrier and replicate in usually non-permissive human cells [76]. This drastic effect is achieved by the introduction of only 4 mutations in the M112/113 gene region. One deletion of 3 amino acids, one silent mutation and 2 point mutations only affecting E1p87 allow the virus to replicate in usually non-permissive RPE-1 cells. These mutations in the E1 proteins led to a higher number of disrupted PML nuclear bodies compared to wild type infected cells. However this increased disruption of PML

nuclear bodies does not seem to be the primary reason for overcoming the species barrier [76].

3.2 Splicing and alternative splicing

Most eukaryotic genes are subject to splicing, a process which involves the removal of parts of the premature messenger RNA (pre-mRNA) called introns. The remaining parts of the pre-mRNA, called exons, are then joined ("splicing"). Splicing is a crucial step in the generation of the mature messenger RNA (mRNA) [91, 92].

Introns are marked at their 5′ and 3′ ends by so called splice sites, with the 5' start of the intron marked by a splice donor site and its 3' end by a splice acceptor site. These splice sites are highly conserved consensus sequences (FIG 5). In most cases, the intron sequence that is removed starts with the nucleotides GU at its 5′ end, and ends with AG at its 3′ end. Additionally, 18 to 40 nucleotides upstream from the 3′ end of an intron the essential branch site can be found. This is a rather poorly conserved region which always contains a critical adenine [93].

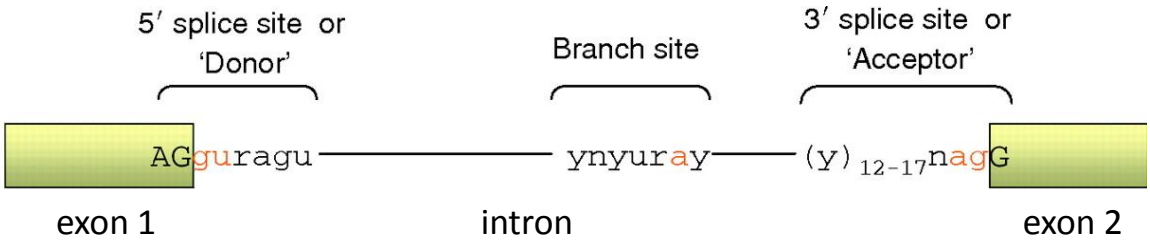


FIG 5. Schematic representation of the consensus sequences for the 5′ splice site (donor), branch site and 3′ splice site (acceptor). Capital letters represent bases in the exon, small letters represent bases in the intron with n being any base, r being A or G and y being C or U. Modified from: Srebrow and Kornblihtt, The connection between splicing and cancer, Journal of Cell Science, 2006.

The actual splicing is a multi-step process which involves several proteins forming a complex machinery termed the spliceosome [94]. Important components of this molecular machinery are several small nuclear ribonucleic proteins (snRNPs). The molecular mechanisms of splicing as well as the structure of the spliceosome will not be described in detail here. In a simplified way, splicing occurs as follows (FIG 6): SnRNP U1 attaches to the splice donor site and snRNP U2 attaches to the branch site of an intron. These splicing factors serve as beacons for the assembly of the spliceosome, recruiting the crucial snRNPs U5 and U4/6. Afterwards the spliceosome brings the two

exons in close proximity leading to a looping out of the intron sequence. This allows a nucleophilic attack of the OH-group of the adenine of the branch site to the splice donor site, leading to the formation of a closed loop. The free OH-group of the splice donor site can then attack the splice acceptor site joining the two exons. The looped out fragment, also called lariat, is then released together with the spliceosome and degraded in the nucleus [93, 95, 96].

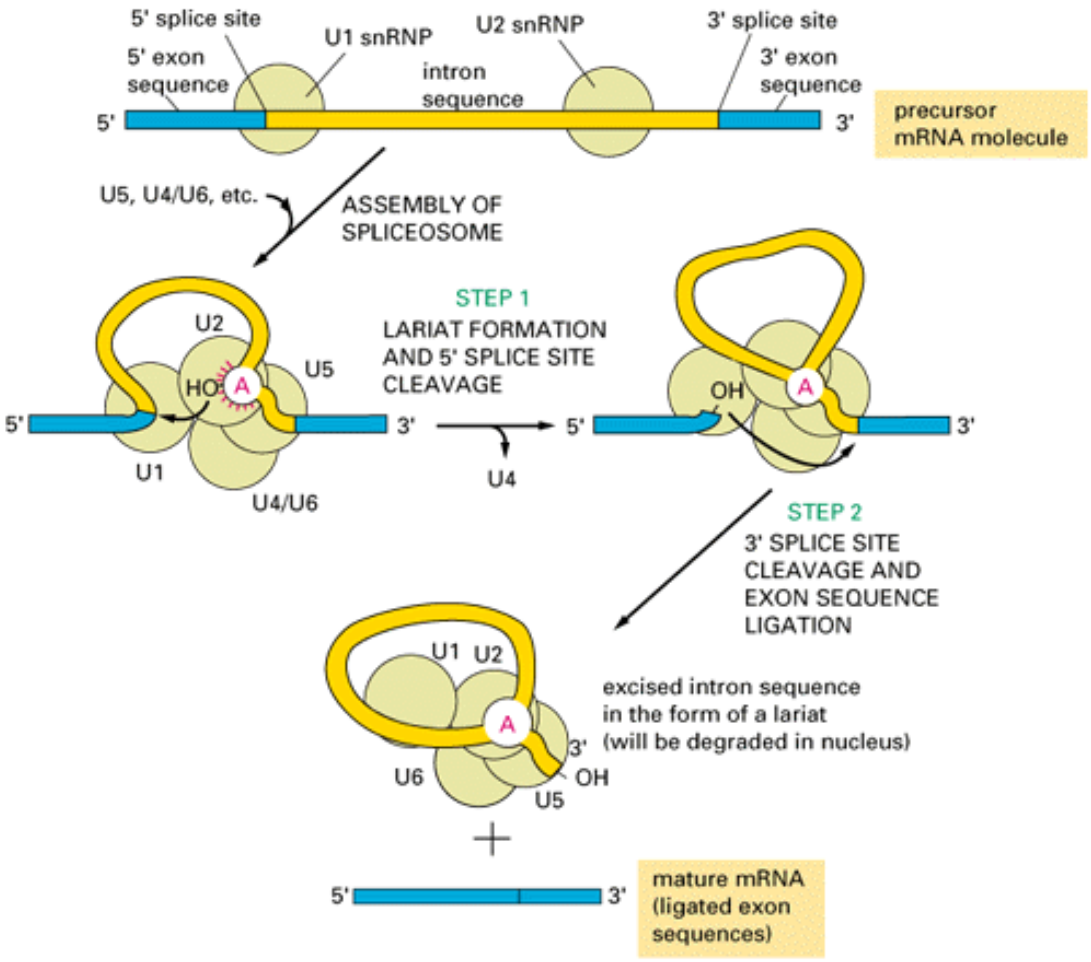


FIG 6. Simplified splicing mechanism. This schematic illustrates the excision of an intron (yellow) flanked by two exons (blue) leading to a spliced RNA product. Taken from: Alberts, Brey, Lewis, Molecular Biology of the Cell, 3rd edition, 1994.

Splicing is also used in eukaryotic cells to express multiple protein isoforms from one gene by the generation of differentially spliced mRNAs. This process is called alternative splicing [97-99]. Alternative splicing is regulated by a highly diverse system of splicing activators and splicing repressors with little conservation between different alternative spliced genes [100]. However, several modes of alternative splicing can be distinguished (FIG 7) [93, 98, 101, 102]. Exons flanked by introns can be spliced out in some

transcripts; a process known as exon skipping. Different spliced mRNAs can also originate from the use of alternative splice donor or splice acceptor sites. It is also possible that introns are retained and their genomic information is translated into amino acids.

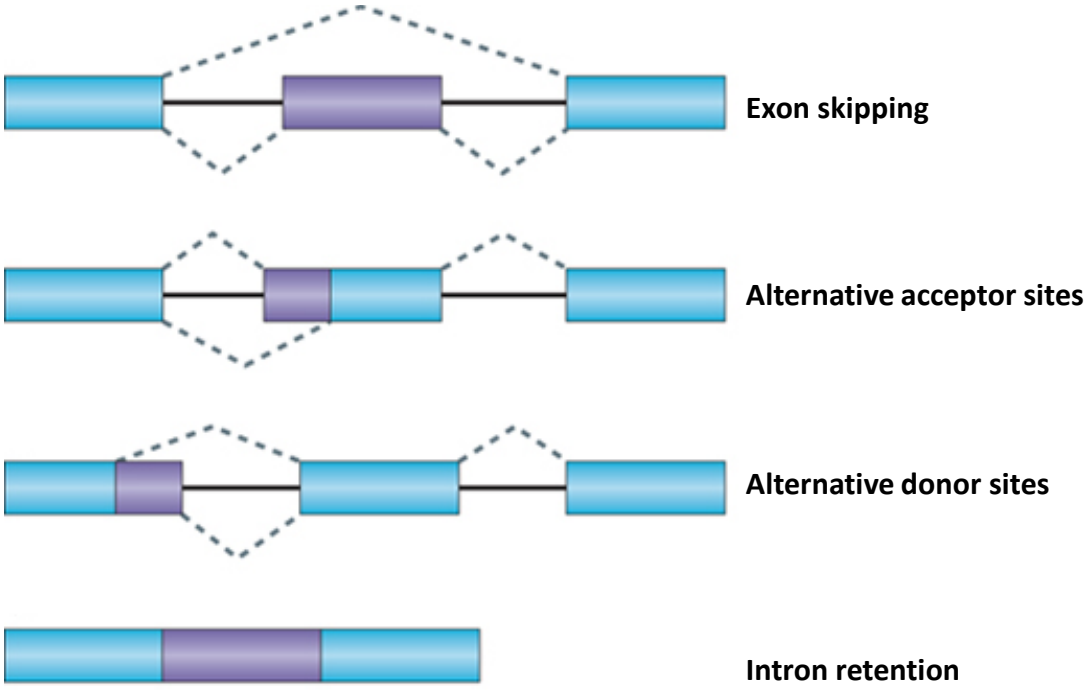


FIG 7. Modes of alternative splicing. Constitutive exons are shown in blue, differentially spliced regions are shown in purple. Introns are shown as lines. Dotted lines indicate different options of splicing. Modified from: Keren et al., Alternative splicing and evolution: diversification, exon definition and function, Nature Reviews Genetics 11, 2010

3.2.1 Alternative spliced cytomegalovirus transcripts

Although more than 90 % of cellular transcripts consist of exons and introns and require splicing to become mature protein-coding mRNAs [99, 103], the majority of herpesvirus proteins are translated from unspliced transcripts. As unspliced transcripts are exported with lower efficiency from the nucleus to the cytosol, herpesviruses express mRNA export factors that increase nuclear export and translation of unspliced viral transcripts [104]. However, a substantial number of herpesviral proteins are expressed from spliced transcripts. Interestingly, many of these are transcribed during latency or at the beginning of the lytic cycle.

It has been known for a long time that HCMV expresses several proteins from spliced

transcripts. In fact, a more recent analysis of the HCMV transcriptome by deep sequencing identified a surprisingly large number of splice junctions that were predicted to affect 58 protein coding genes, indicating that splicing of viral transcripts is more common than previously recognized [105]. The number and diversity of viral proteins is further extended by alternative splicing, which allows a single gene to code for several protein products.

Besides the UL112/113 gene locus, which this works focuses on, only 4 HCMV ORFs have been reported to be differentially spliced (FIG 8) [105, 106]. Among them is the very well-studied major immediate early locus UL122/UL123 which codes for several isoforms of the immediate early 1 (IE1) and 2 (IE2) proteins [107-111]. Alternative splicing is also found in the UL36-UL38 gene region which codes for proteins involved in apoptosis regulation as well as viral DNA replication [112-116]. The HCMV US3 gene gives rise to several transcripts by alternative splicing and the resulting protein products seem to be involved in immune modulation [113, 117-119]. Alternative splicing also occurs in the UL73-UL74/UL75 intergenic region [120].

Compared to HCMV, alternative splicing is only very poorly studied in MCMV. Nevertheless most of the genes alternatively spliced in HCMV also show alternative splicing in MCMV [121]. Due to that conservation it can be assumed that alternative splicing is important for the function of those viral genes and overall viral replication.

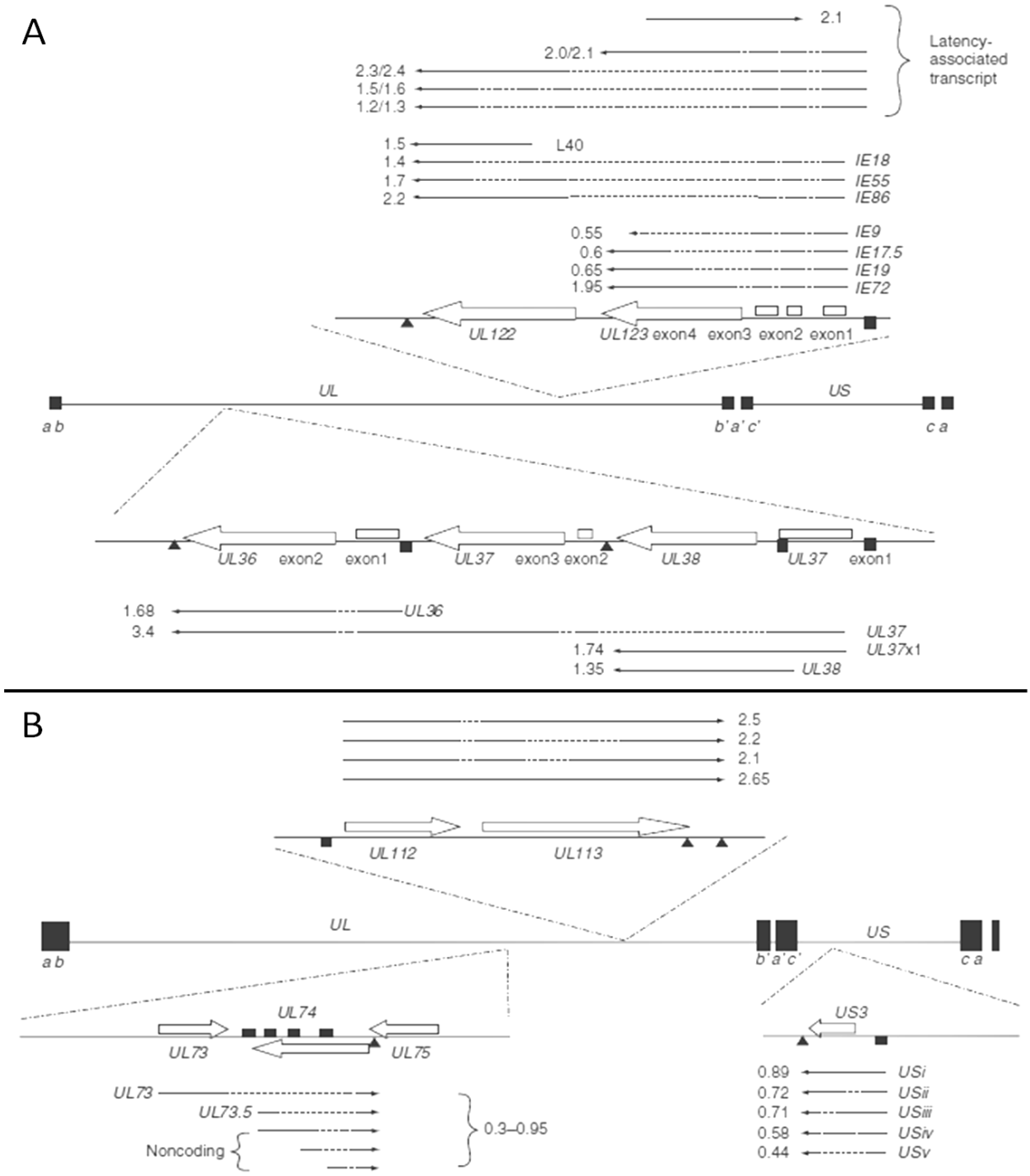


FIG 8. Alternative spliced transcripts of HCMV. Transcripts resulting from alternative splicing of (A) the UL122–UL123 and UL36–UL38 gene regions, as well as (B) the UL112–UL113, UL72 and US3 gene regions. Names or lengths (kb) of the transcripts are labeled alongside. A black square represents a TATA box; a black triangle represents a poly-A signal. Modified from: Ma et al., Human CMV transcripts: an overview. Future Microbiology 2012.

4 Aims of the study

Compared to most other viruses, herpesviruses have very large and complex genomes, upon which they encode a large number of viral proteins. This complexity is further increased by the fact that a subset of viral transcripts is subject to splicing or even alternative splicing. Recent viral transcriptome analyses by RNA-seq have demonstrated that splicing of herpesviral transcripts is more common than previously thought [105, 122-125].

Among those spliced viral transcripts are the CMV Early 1 transcripts coded by the MCMV gene region M112/113 and its HCMV homolog UL112/113. They encode the E1 proteins, which have been shown to exist in at least 4 isoforms, translated from alternatively spliced transcripts in both MCMV and HCMV [66-68] and seem to play a crucial role in viral gene expression and DNA replication [69, 70, 77, 78, 81, 84, 85, 87, 126]. The E1 proteins can be described as the archetype of the class of CMV early proteins although little is known about them. Therefore, the aim of this study was to further characterize this prominent set of viral proteins.

It can be assumed that the four E1 protein isoforms do not have entirely redundant functions. Consequently there was the need for a strategy to investigate the function(s) of single protein isoforms of the alternatively spliced gene regions. One major aim of this study was to inactivate the E1 isoforms individually by using splice donor and acceptor mutagenesis, intron deletion and substitution mutagenesis, and the introduction of stop codons (nonsense mutagenesis).

In addition, virus mutagenesis was combined with SILAC, affinity purification and LC-MS/MS to identify interaction partners of the large E1 isoform of MCMV. The potential cellular interaction partners were used to help decipher the function of the largest MCMV E1 isoform.

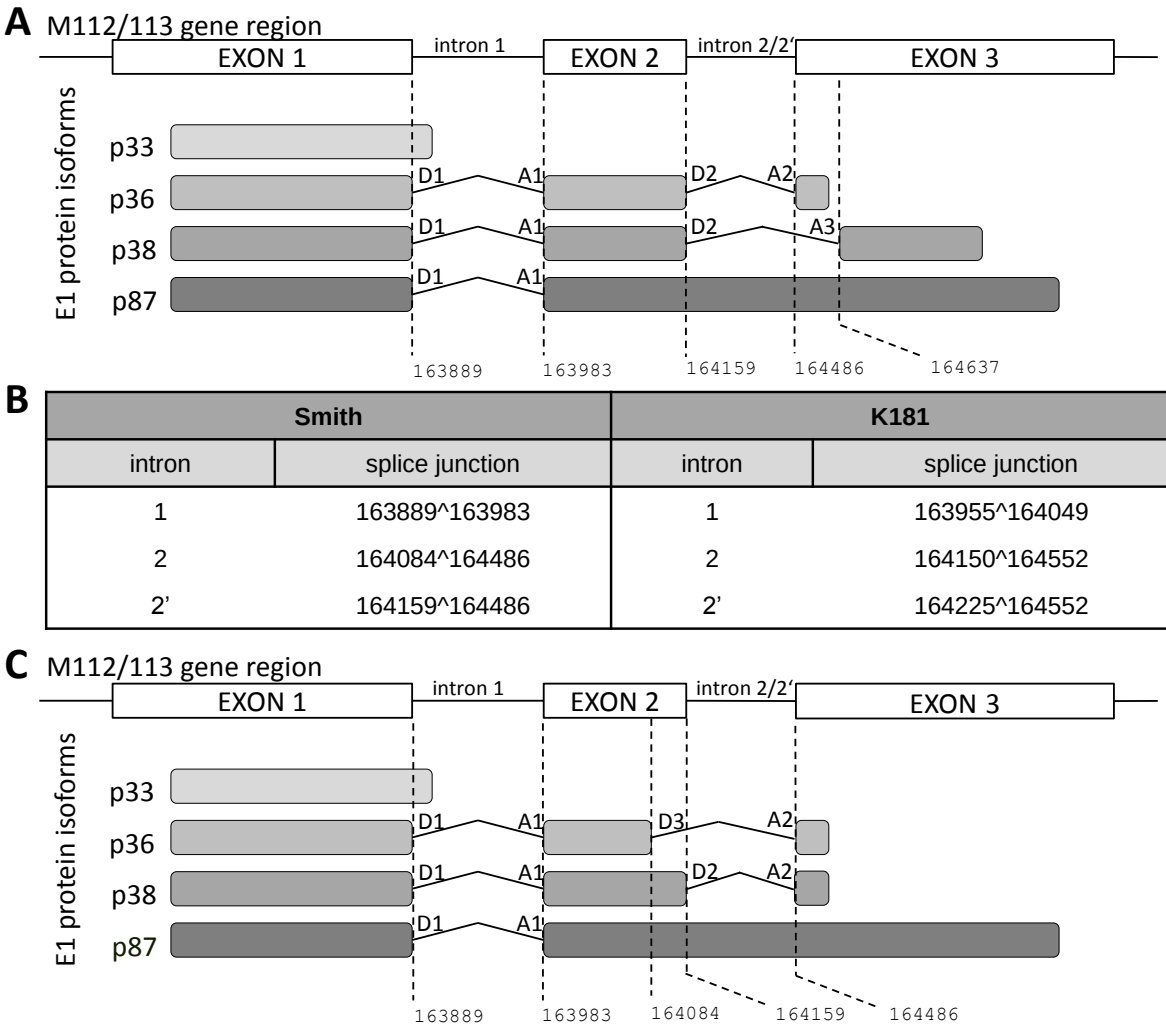
Taken together these strategies are an opportunity to gain new insights into the individual function(s) of single isoforms of the alternatively spliced M112/113 and UL112/113 gene regions.

5 Results

5.1 RNA-seq analysis of the MCMV M112/113 gene region leads to a revised M112/113 splicing pattern.

Previous studies have shown that the M112/113 region consists of 3 exons interspaced by 2 introns and encodes 4 proteins derived from differentially spliced transcripts [66, 68]. The small E1p33 isoform is described to be translated from an unspliced transcript, whereas the large E1p87 isoform is derived from a transcript where only the first intron is spliced out. The two medium sized isoforms E1p36 and E1p38 are translated from fully spliced transcripts and it has been reported that they make use of two different splice (A2 and A3) acceptors, using a shared splice donor (D2). However the exact genomic position of the splice acceptor site giving rise to the E1p38 isoform has never been described. Therefore, the position of this splice acceptor was determined using Hbond score [127]. This algorithm predicted the E1p38 acceptor at position 164637 (FIG 9A).

To verify the previously predicted splicing pattern and determine the relative abundance of the individual M112/113 transcripts in MCMV-infected NIH-3T3 fibroblasts, whole transcriptome shotgun sequencing (also called RNA-seq) was applied. This was done to identify the splice junctions of viral transcripts. The experiments were performed by Daniela Indenbirken (technology platform Next Generation Sequencing, HPI, Hamburg, Germany) and the resulting raw data were analyzed by Malik Alawi (Bioinformatics Core at the University Medical Center Hamburg-Eppendorf, Germany). The identified splice junctions largely confirmed the previously predicted splicing pattern with two notable exceptions: the predicted splice acceptor site A3 was not identified, but instead a new splice donor site D3 at nucleotide position 164084 (FIG 9B). This leads to a revised splicing pattern of the M112/113 region in the MCMV Smith strain (FIG 9C). The same splice junctions were identified by RNA-seq of NIH-3T3 cells infected with the MCMV K181 strain, indicating that the revised splicing pattern is not unique to the MCMV Smith strain.



Smith		K181	
intron	splice junction	intron	splice junction
1	163889^163983	1	163955^164049
2	164084^164486	2	164150^164552
2'	164159^164486	2'	164225^164552

FIG 9. Analysis of the intron-exon structure of the MCMV M112/113 gene. (A) Schematic representation of the M112/113 locus and the four E1 proteins according to Ciocco-Schmitt et al. Splice donor and acceptor sites are marked with D and A, respectively. Nucleotide positions of exon boundaries are indicated. (B) Splice junctions were detected by RNA-Seq of cells infected with MCMV Smith or MCMV K181. (C) Schematic representation of the revised M112/113 locus based on RNA-Seq data.

In addition, the relative abundance of spliced and unspliced transcripts was analyzed at different times post infection. The ratio of spliced to unspliced transcripts for those spanning intron 1 did not change significantly over the course of infection. The relative abundance of unspliced transcripts retaining intron 2 also remained largely unchanged, however the relative abundance of D2^A2 and D3^A2 spliced transcripts changed over time (FIG 10). The D2^A2 and D3^A2 junctions being/were equally abundant at 3 hours post infection, but during the course of infection the number of transcripts containing a D3^A2 junction slightly increased, consequently leading to a slight decrease in transcripts containing a D2^A2 junction. The overall percentage of unspliced transcripts

was very low (less than 3%) for both introns, indicating low amounts of the isoforms E1p33 and E1p87 compared to the two medium sized E1 isoforms. Other splice junctions using different donor and/or acceptor sites or other combinations of the described donor and acceptor sites have also been detected. Most of these splice junctions were not conserved between MCMV Smith and K181 and were detected in only a few sequence reads. Hence these low-abundant splices were not further investigated.

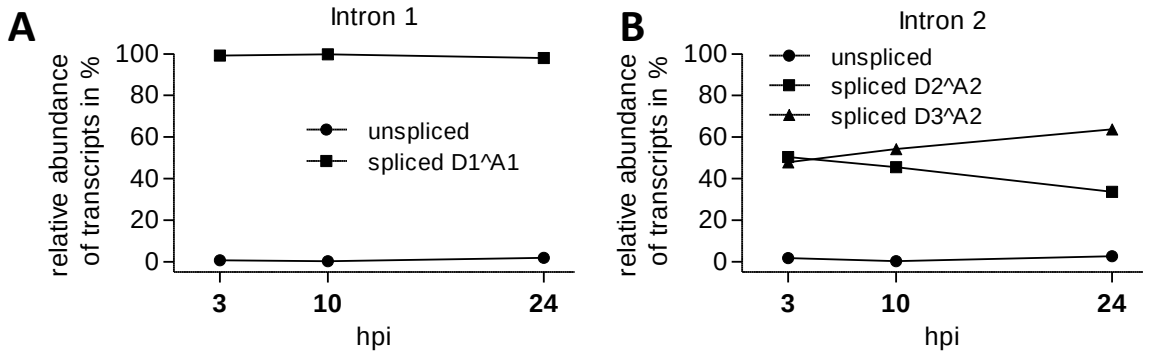


FIG 10. Relative abundance of spliced and unspliced transcripts for the M112/113 intron 1 (A) and intron 2 (B).

5.2 The revised M112/113 splicing pattern was verified by plasmid mutagenesis.

In order to verify the splicing pattern predicted by the RNA-seq data the M112/113 coding region was mutated by *en passant* BAC mutagenesis in a way to inactivate individual E1 isoforms [128]. Mutant and wt M112/113 gene regions were subsequently cloned in the plasmid expression vector pcDNA3 (FIG 11). The following mutations were introduced to the inactive E1 isoforms: Intron 1 was removed in the Δp33 mutant, leading to a fusion of exons 1 and 2. In the Δp36 mutant the splice donor D3 was mutated by two nucleotide exchanges that do not affect the amino acid sequence encoded by the unspliced transcript. In the same manner, splice donor D2 was mutated to generate the Δp38 mutant. A single nucleotide mutation was introduced in the splice acceptor A2 in order to eliminate expression of both medium-sized isoforms (mutant Δp36Δp38). Unfortunately, there was no way to completely eliminate the large p87 isoform without affecting other isoforms. Hence, two point mutations were introduced shortly after splice donor D2. These point mutations lead to two premature stop codons,

which should result in a MCMV mutant containing a truncated p87 protein lacking the last 292 amino acids.

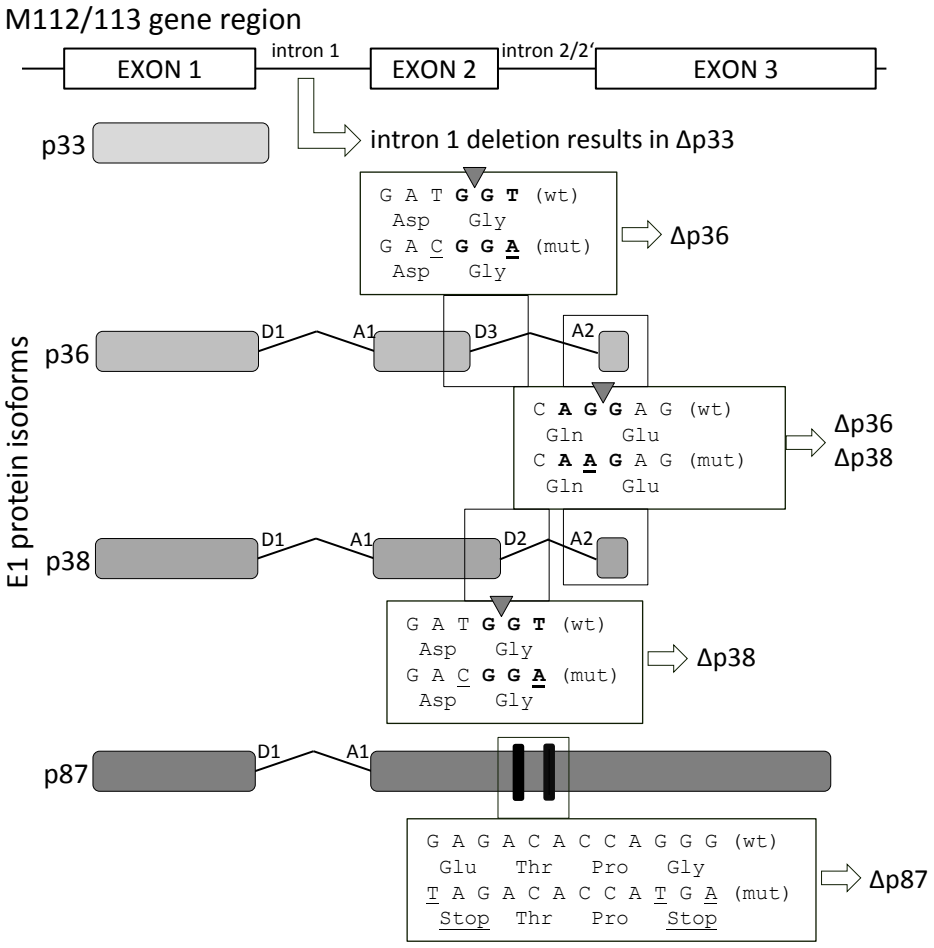


FIG 11. Mutagenesis of the M112/113 locus. Schematic view of the MCMV M112/113 locus and the four proteins expressed from this locus by alternative splicing. Mutations introduced to inactivate the expression of individual protein isoforms are shown. Splice donor and acceptor sites are indicated by arrow heads, and the minimal consensus sequences are shown in bold. Nucleotide and amino acid changes are underlined.

To determine whether the mutations lead to a loss of individual E1 protein isoforms, the mutated plasmids were transfected into NIH-3T3 cells and their cell lysates were analyzed by WB using an E1-specific antiserum recognizing all four E1 isoforms [68]. As expected, cells transfected with the M112/113 Δp33, Δp36, and Δp87 plasmids expressed all E1 protein isoforms except for those intentionally deleted (FIG 12). NIH-3T3 cells transfected with the M112/113 Δp38 or Δp36Δp38 neither expressed the deleted isoforms nor the p33 isoform (FIG 12A). However, a weak expression was detected by immunoblot in transfected 293A cells (FIG 12B), suggesting that p33 expression from these two plasmids was weak but not absent. It is also worth noting

that a truncated p87 protein was not observed in cells transfected with the Δp87 plasmid (Fig. 2B). However, since the predicted size of the truncated p87 protein is 36 kDa and similar to that of the p36 and p38 isoform, it is possible that a band representing the truncated p87 protein is hidden within the bands representing the p36 or p38 isoforms. In summary, the results of the M112/113 plasmid mutagenesis confirm the splicing pattern prediction based on the RNA-seq data.

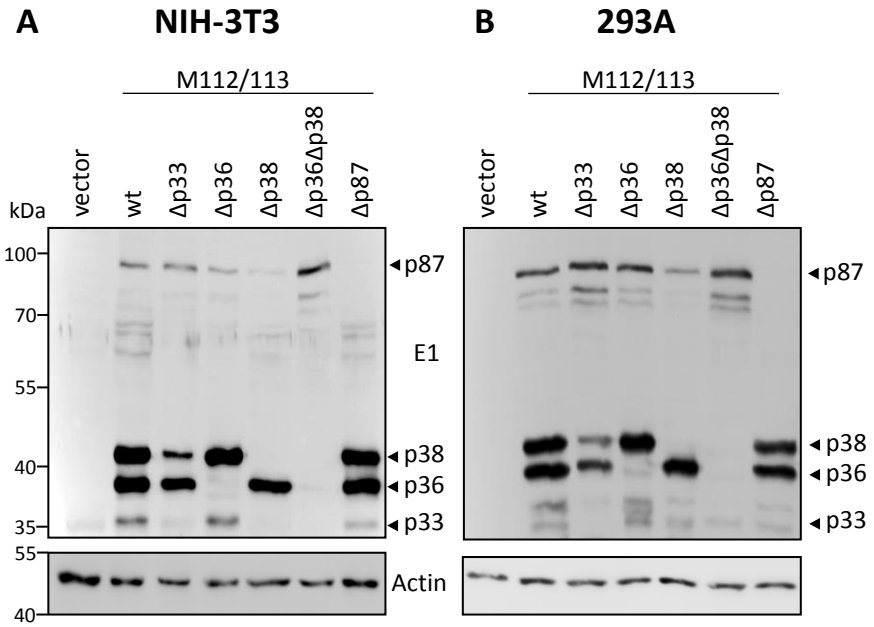


FIG 12. MCMV E1 isoform expression from E1 inactivation plasmids. (A) NIH-3T3 or (B) 293A cells were transfected with empty vector or expression plasmids containing the wt or mutant M112/113 sequence. E1 protein expression was detected by immunoblotting using an E1-specific antiserum. Actin was detected as a loading control.

5.3 Individual inactivation of the E1 isoforms p36, p38 and p87 has different impact on MCMV replication.

To investigate whether the loss of E1 isoforms influences the ability of MCMV to replicate in cell culture, MCMV mutants lacking individual E1 isoforms were constructed. The same mutations as described for the M112/113 plasmids (FIG 11) were introduced into a BAC clone of MCMV-GFP, i.e. wt MCMV (Smith strain) expressing GFP. The MCMV-GFP backbone was chosen for easier detection of poorly replicating virus mutants after BAC transfection. In addition an E1 stop mutant was constructed in which codons 6, 8 and 9 of M112 were changed to stop codons. Mutant BACs were analyzed by restriction digestion, gel electrophoresis and sequencing of the M112/113 region (FIG 13). As

controls for MCMV M112/113 mutants who showed a replication defect, revertant MCMV mutants in which the M112/113 mutations were reverted to the wt sequence were constructed. The revertant viruses allow to control that the observed replication defects of the mutant viruses is a result of the introduced mutations rather than an accidental mutation elsewhere in the viral genome. The mutant and revertant MCMV BACs were then transfected into NIH-3T3 fibroblasts to reconstitute mutant viruses.

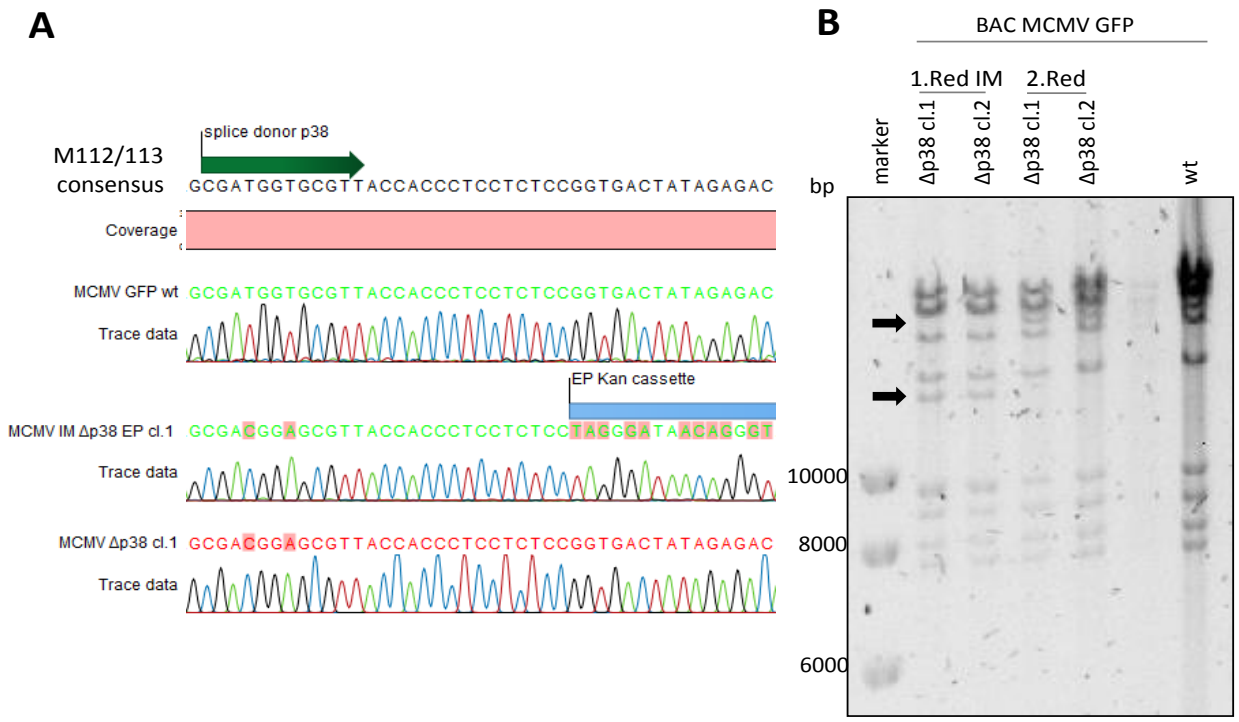


FIG 13. Control of BAC mutagenesis. The MCMV Δp38 BAC is shown exemplary. (A) Sequencing of MCMV Δp38 BAC clones. Electropherograms of MCMV GFP, MCMV Δp38 IM and MCMV Δp38 BACs are shown. The p38 donor site and the EP Kan cassette are annotated. Point mutations are highlighted. (B) HindIII restriction pattern of MCMV BAC DNA from first (1. Red IM) and second (2. Red) recombination step. MCMV GFP BAC (wt) is used as control. Arrows indicate predicted band pattern differences.

In spite of numerous attempts, the E1 stop mutant could not be reconstituted. However, the revertant virus, in which the stop mutations had been changed back to the original sequence, replicated like the wt virus (data not shown). This finding is consistent with a previous publication reporting that an MCMV M112/113 deletion mutant was unable to replicate in cultured fibroblasts [76] and suggests that the E1 proteins are essential for MCMV replication. However, the possibility that individual E1 isoforms are nonessential remained to be investigated.

Viral replication and spread was not detectable upon transfection of BAC DNA of the MCMV Δp87 mutant into NIH-3T3 cells. However, the corresponding revertant virus (Rev p87) replicated with wt kinetics (FIG 15A) suggesting an essential role of the p87 isoform for MCMV replication.

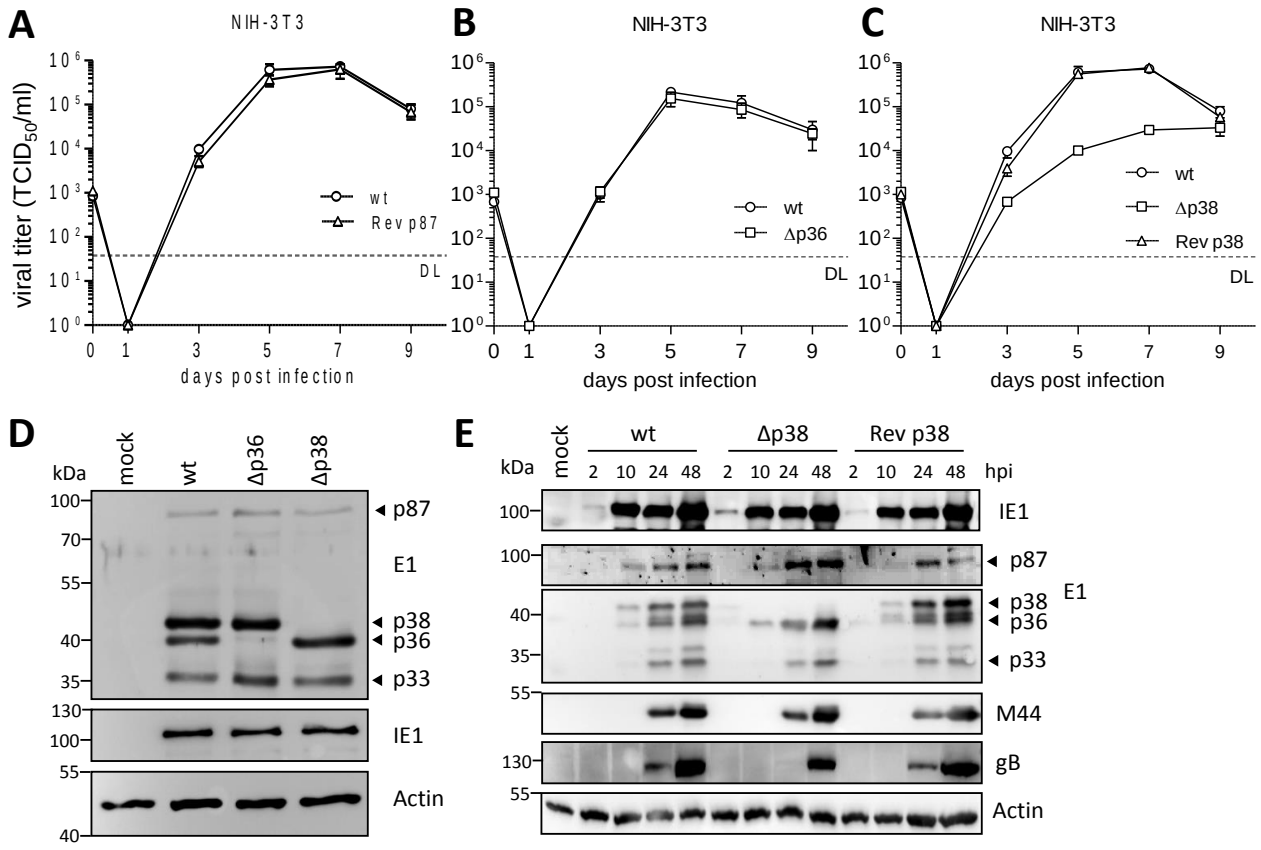


FIG 14. Replication kinetics and E1 protein expression of MCMV M112/113 mutants. (A) NIH-3T3 cells were infected at an MOI of 0.02 $TCID_{50}$/cell with wt MCMV and Rev p87. Virus released into the supernatant was quantified by titration. Replication kinetics of (B) wt and Δp36 and (C) wt, Δp38, and Rev p38 were determined as described above. Means ± SEM of experiments done in triplicate are shown. DL, detection limit. (D) NIH-3T3 cells were infected at an MOI of 3 $TCID_{50}$/cell with wt and mutant MCMVs. E1 protein expression was determined by immunoblotting. IE1 was used as infection control and β-actin as loading control. (E) NIH-3T3 cells were infected with MCMV wt, Δp38 or Rev p38 at an MOI of 0.5 $TCID_{50}$/cell. Cells lysates were harvested at different times post infection. Expression levels of the viral proteins IE1, E1, M44, and gB were analyzed by immunoblot. The upper part of the E1 immunoblot is shown as a longer exposure for better visualization of the p87 isoform.

In order to rescue this phenotype I tried to complement the loss of E1p87 in *trans*. To do that, a C-terminally hemagglutinin (HA) tagged E1p87 ORF containing retrovirus was constructed and NIH-3T3 were transduced using this retrovirus. Approximately 4 days post transduction MCMV Δp87 BACs were transfected in those cells. The expression of E1p87-HA was clearly detectable (FIG 14) in transduced cells and MCMV wt BAC could easily be reconstituted in those cells. However the expression of E1p87-HA in *trans* did not rescue the growth defect of the MCMV Δp87 mutant.

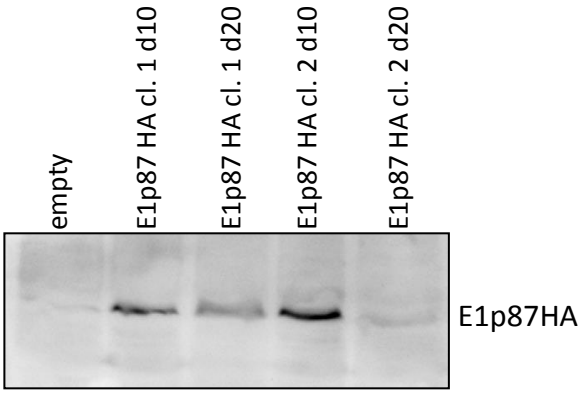


FIG 15. MCMV E1p87-HA cell line. NIH-3T3 cells were transduced using a retrovirus carrying an E1p87-HA ORF. E1p87 expression was determined by immunoblot using a HA-specific antibody. Cells were lysed 10 and 20 days post transfection respectively.

The MCMV Δp36 and Δp38 mutants were easily reconstituted and replicated well in NIH-3T3 fibroblasts. While the MCMV Δp36 mutant replicated with wt kinetics in NIH-3T3 cells (FIG 15B), the Δp38 mutant replicated to titers 10 to 100-fold lower than the parental virus. The corresponding revertant (Rev p38) replicated with wt kinetics (FIG 15C). Immunoblot analysis of cells infected with the Δp36 and Δp38 mutants showed the absence of the deleted E1 isoforms while the remaining isoforms were expressed at similar levels as in cells infected with wt MCMV (FIG 15D).

To further investigate the phenotype of the Δp38 mutant, I analyzed viral protein expression up to 48 hrs after low-MOI infection. In fibroblasts infected with the Δp38 mutant, expression of the viral IE1, E1, and M44 proteins was similar to the one observed wt or Rev p38-infected cells, but expression of the late protein gB was delayed (FIG. 15E). Since late gene expression occurs only after viral DNA replication, the delayed gB expression indicates impaired viral DNA replication and/or reduced late gene expression

Taken together these results suggest an essential role for E1p87 for MCMV replication in fibroblasts. In fibroblasts E1 p38 seems to play an augmenting role in viral replication, whereas E1p36 is dispensable for MCMV replication. Interestingly the intensity of the E1 bands in the immunoblots of MCMV wt infected cells reflects the differences in the abundance of E1 transcripts (FIG 10). There is less expression of E1p33 and E1p87 compared to the two medium sized isoforms which displays a direct correlation between their transcripts and protein isoforms.

5.4 Simultaneous inactivation of E1p36 and E1p38 is rescued by spontaneous mutation.

As the E1 p36 and p38 isoforms were not essential for MCMV replication, it was worth investigating whether a virus lacking both of these isoforms could replicate. As shown in the plasmid mutagenesis experiment, mutation of the A2 splice acceptor site (FIG 11) resulted in a loss of the p36 and p38 isoforms. When a mutant MCMV BAC carrying the same mutation (Δp36Δp38) was transfected into NIH-3T3 cells, the replicating virus could not be recovered despite several attempts. In contrast, a revertant (Rev p36p38) replicated with the same kinetics as the parental wt virus (FIG 16A). Only on one occasion a single plaque was observed after transfection of the MCMV Δp36Δp38 BAC, suggesting that a compensatory mutation might have occurred during this instance. This spontaneously mutated virus could be propagated and further analyzed.

The absence of both p36 and p38 isoforms was confirmed following analysis of lysates of infected cells by immunoblot. Interestingly, the p33 and p87 isoforms were expressed at higher levels than in cells infected with wt MCMV, and an unexpected additional protein with an apparent molecular mass of approx. 60 kDa was detected by the E1-specific antiserum (FIG. 16B). To investigate the origin of this additional protein, the M112/113 gene of this mutant (named Δp36Δp38mut) was sequenced. In addition to the splice acceptor mutation, a G-to-T mutation was detected at position 164420. This mutation was apparently not present in all viral genomes as the sequencing electropherogram showed two peaks (G and T) of similar amplitude at this position (FIG 16C). This result suggested a mixed infection with Δp36Δp38 and a mutant thereof carrying a G-to-T mutation, which leads to a premature stop and expression of a truncated p87 protein. This hypothesis was confirmed when a part of the M112/113 gene was PCR-amplified and subsequently cloned in a plasmid vector. Following this, 9 plasmid clones were sequenced: 5 clones carried the wt sequence and 4 carried the G-to-T mutation,

suggesting that the two mutant viruses were present in similar amounts. Attempts to separate the two mutants by limiting dilution failed, suggesting that the two mutants complement each other and neither can replicate alone. However, the mix of mutants (Δp36Δp38mut) replicated to slightly lower titers than the wt virus (FIG 16D).

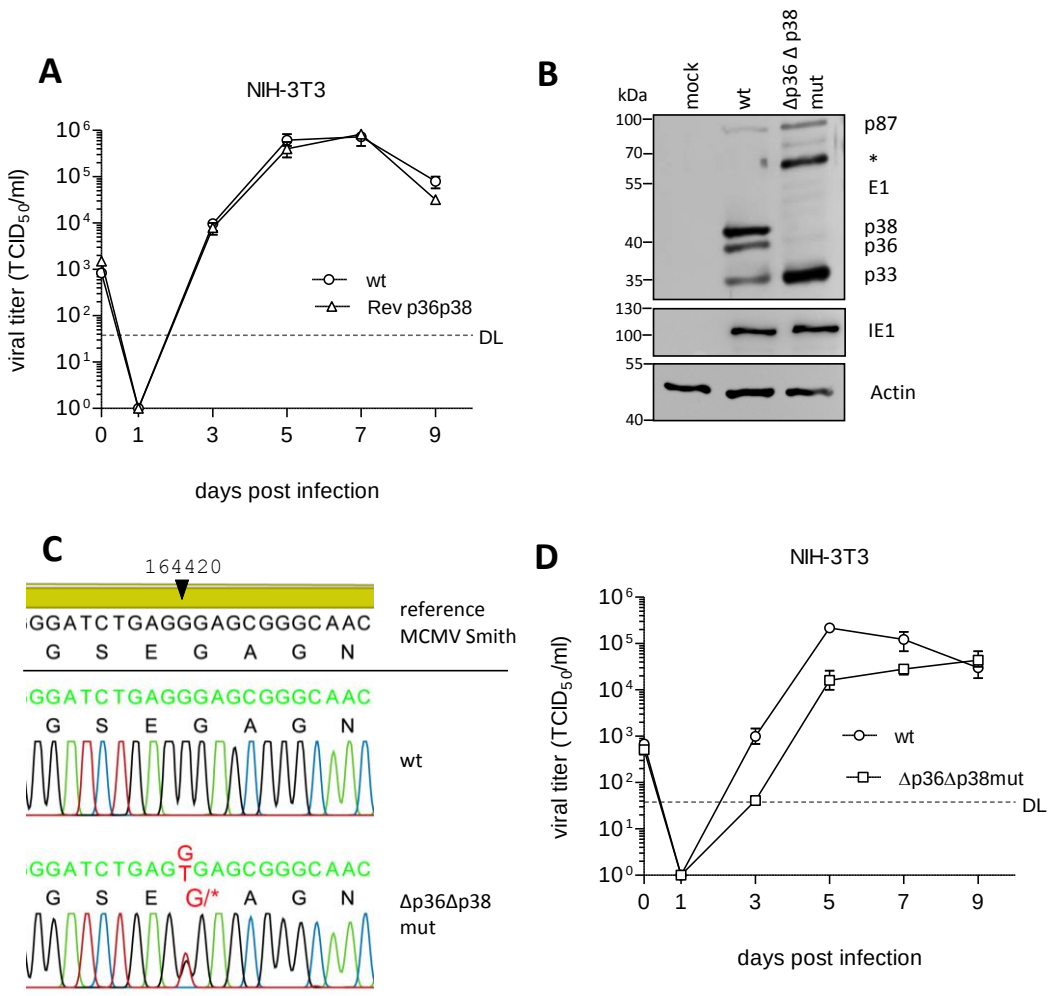


FIG 16. Characterization of a spontaneously adapted Δp36Δp38 mutant. (A) NIH-3T3 cells were infected at an MOI of 0.02 $TCID_{50}$/cell with wt MCMV or Rev p36p38. Viral replication kinetics were determined as described in the previous figure. Means ± SEM of experiments done in triplicate are shown. DL, detection limit. (B) NIH-3T3 cells were infected at an MOI of 3 $TCID_{50}$/cell with wt MCMV or an MCMV mutant that spontaneously emerged in fibroblasts transfected with the MCMV Δp36Δp38 BAC (Δp36Δp38 mut). E1 protein expression was analyzed by immunoblotting. An asterisk marks a new E1 isoform expressed by the spontaneously occurring MCMV mutant. (C) Identification of a mutation by sequencing the exon 2 to exon 3 region. The nucleotide position of the mutation is given above. (D) Viral replication kinetics of wt MCMV and Δp36Δp38 mut were determined as described in the previous figures.

To rule out the possibility that a compensatory second-site mutation rather than the G-to-T mutation within M112/113 was responsible for the replication of Δp36Δp38mut, I introduced the G164420T mutation into Δp36Δp38 by BAC mutagenesis. This mutant, Δp36Δp38/p60, should express only p33 and the truncated p60 protein, but not p36, p38, or p87. The Δp36Δp38 and Δp36Δp38/p60 BACs were transfected into NIH-3T3 cells, separately and in combination. As expected, neither the Δp36p38 nor the Δp36Δp38/p60 BAC alone was able to produce replicating virus. However, co-transfection of both BACs resulted in plaque formation (FIG 17). This result confirmed that the G164420T mutation leading to a truncated protein (p60) is sufficient to restore virus replication.

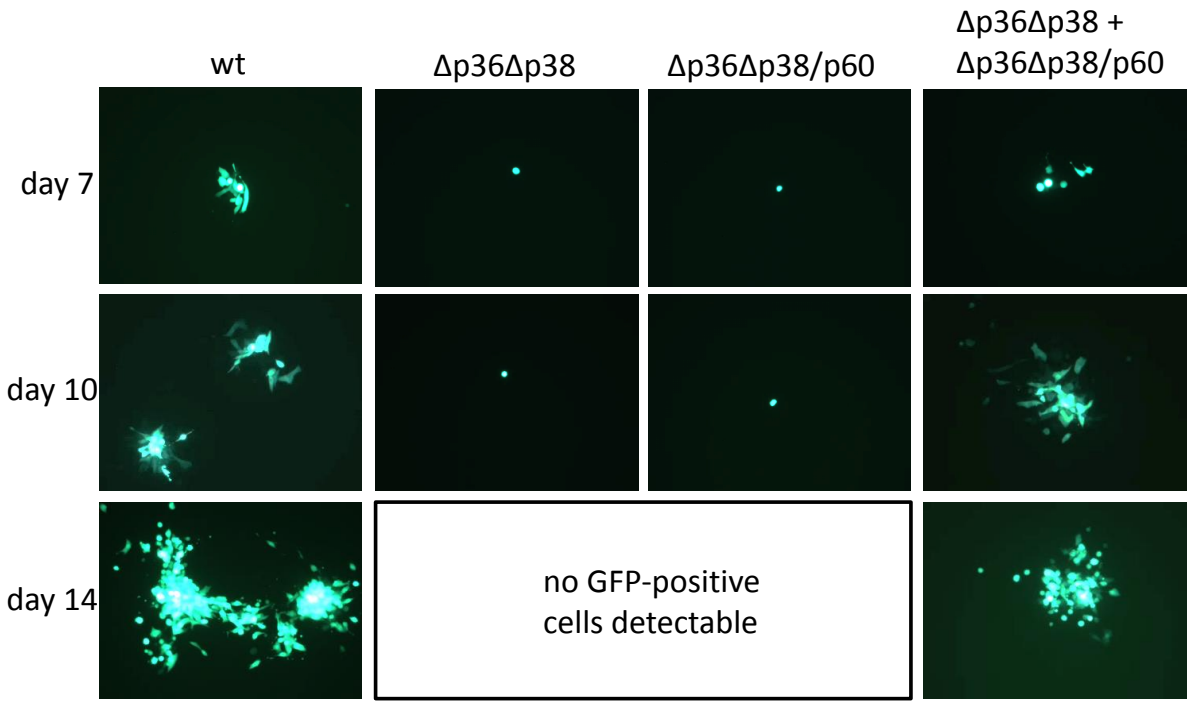


FIG 17. Co-transfection of Δp36Δp38 and Δp36Δp38/p60 BACs. Cells were transfected with MCMV wt BAC, Δp36Δp38, Δp36Δp38/p60 or a combination of the latter two. Formation of viral foci was monitored. Cells were passaged on day 8.

Taken together the results obtained with the MCMV Δp36Δp38 mutants indicate that the loss of one isoform, p36 or p38, but not both, is compatible with MCMV replication. Interestingly a mix of viral quasispecies seems to be able to compensate the individual replication restrictions of single virus mutants, at least in the case of inactivated E1 proteins.

5.5 The presence of the M112/113 intron 1 is necessary for MCMV replication.

Surprisingly, the MCMV Δp33 mutant (which lacks intron 1) could not be reconstituted. This result was unexpected as the p33 protein is almost identical to the amino terminal part of p36, p38, and p87 and contains only one lysine residue at its C-terminus that is not present in the other E1 isoforms. Therefore the role of the p33 isoform and intron 1 was investigated in more detail.

First I tested whether the E1p33 protein was required for MCMV replication. To do this, the E1p33 ORF driven by a PGK promoter was inserted into the nonessential m02 to m06 region of the MCMV Δp33 BAC (FIG 18A). This strategy of *cis*-complementation has been reported to be successful for other MCMV mutants [129-131]. However, transfection of the mutant BAC did not result in virus reconstitution, suggesting that ectopic expression of E1p33 is not sufficient to rescue the MCMV Δp33 mutant.

Next I tested whether the presence of M112/113 intron 1 is required for MCMV replication. To do this, a synthetic intron derived from the commonly used plasmid expression vector pCI-Neo (Promega) was inserted into the MCMV Δp33 BAC at the position of the authentic intron 1 (FIG 18B). The synthetic intron 1 (SynI1) is slightly larger than the authentic intron 1 and has a completely different sequence. An unspliced transcript containing SynI1 would not code for p33, but a larger protein with 34 additional amino acids and a predicted mass similar to p36 or p38.

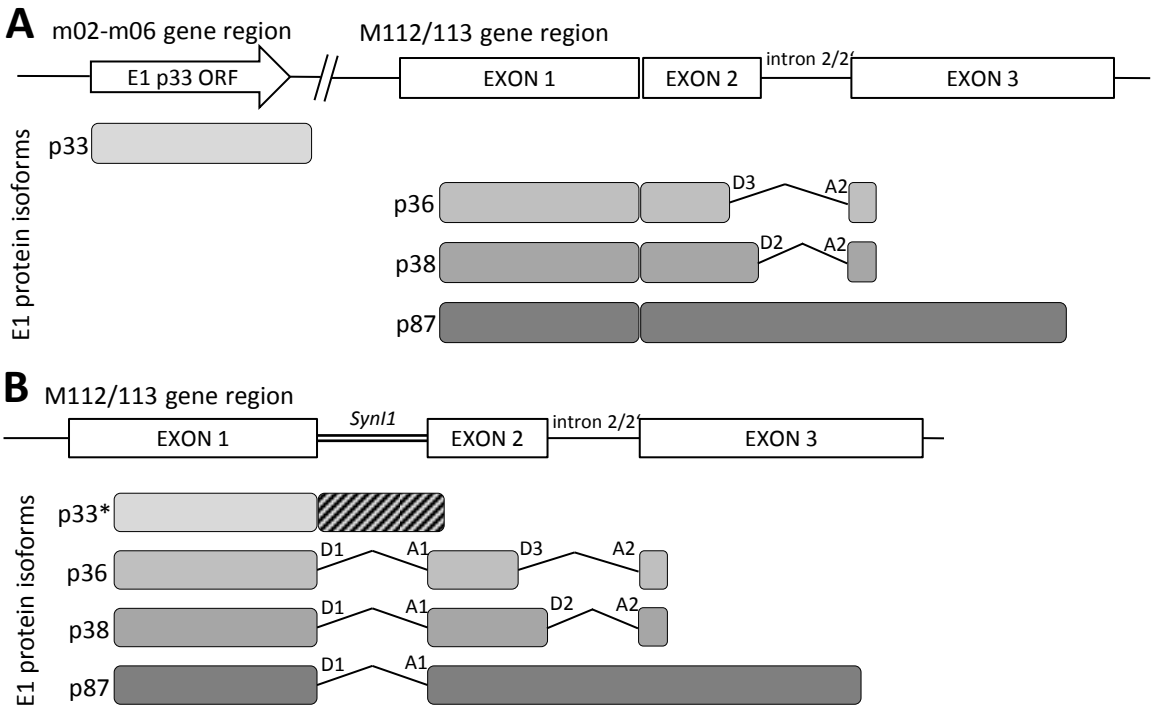


FIG 18. Deletion and substitution mutagenesis of the M112/113 intron 1. (A) Schematic showing the deletion of the M112/1113 intron 1 and insertion of the p33-coding ORF into the nonessential m02-m06 region of the MCMV genome. (B) Schematic of the MCMV SynI1 mutant carrying a synthetic intron in place of the authentic intron 1. An unspliced transcript is predicted to encode an E1 protein significantly larger than p33 (termed p33*).

Strikingly, transfection of the MCMV SynI1 BAC resulted in the recovery of a mutant virus that replicated with wt kinetics in NIH-3T3 fibroblasts (FIG 19A). As predicted, the MCMV SynI1 virus did not express p33 (FIG 19B). However, a new E1 protein product encoded by exon 1 and intron 1 could not be detected, suggesting that such a protein product is either not present or hidden within the band representing the p36 or the p38 isoform. To discern between these two possibilities oligo-dT-primed cDNA from NIH-3T3 cells infected either with wt MCMV or the SynI1 mutant was prepared. Fragments of the M112/113 transcripts were PCR amplified with a primer set binding within exons 1 and 2 and flanking intron 1. The same PCR reaction was conducted using the wt MCMV-GFP BAC (containing the authentic intron 1), the MCMV SynI1 BAC (containing the synthetic intron), and the MCMV Δp33 BAC (lacking intron 1) as templates. As shown in FIG 19C, wt MCMV infected cells contained both spliced and unspliced transcripts. In contrast, only spliced transcripts were detected in cells infected with the MCMV SynI1 mutant, indicating that the synthetic intron is removed with much greater efficiency

than the authentic intron 1. Hence the predicted p33* protein resulting from a read-through into the synthetic intron (FIG 18B) is probably made at very low levels or not at all. Taken together these results strongly suggest that presence of an intron 1 rather than the expression of the E1 p33 isoform is essential for MCMV replication.

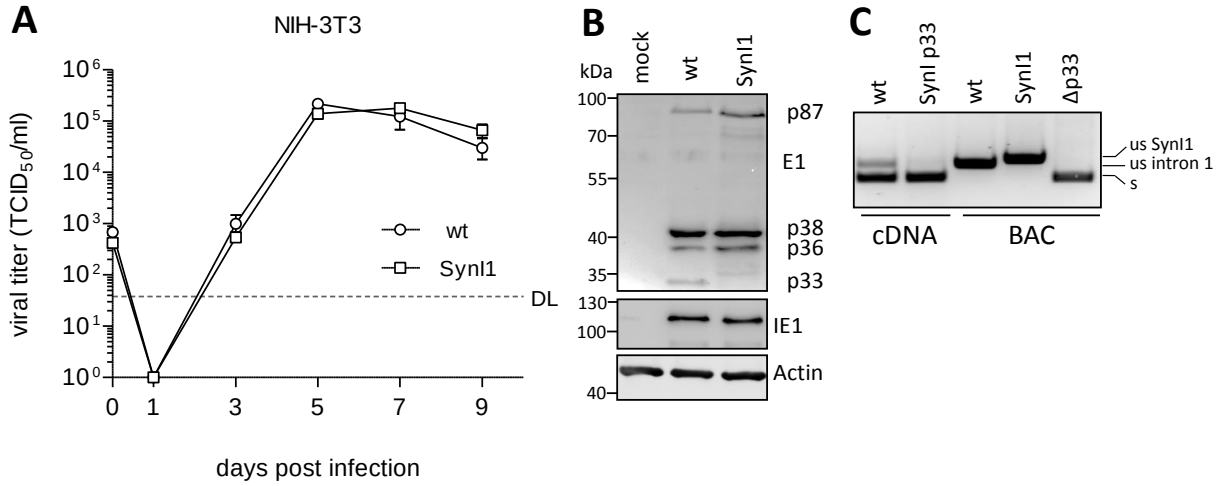


FIG 19. Analysis of MCMV SynI1. (A) NIH3T3 cells were infected at an MOI of 0.02 $TCID_{50}$/cell with wt MCMV and the SynI1 mutant. Virus release into the supernatant was determined by titration. Means ± SEM of experiments done in triplicate are shown. DL, detection limit. (B) NIH-3T3 cells were infected at an MOI of 3 $TCID_{50}$/cell with wt MCMV or the SynI1 mutant. E1 protein expression was analyzed by immunoblotting. IE1 was used as infection control and β-actin as loading control. (C) An M112/113 fragment was PCR-amplified from cDNA of wt MCMV or SynI1-infected cells. The same PCR amplification was conducted using different MCMV BACs as templates. The sizes of unspliced (us) and spliced (s) transcripts are shown.

5.6 Influence of individual E1 isoforms on MCMV replication varies in different cell types.

In NIH-3T3 fibroblasts, the MCMV Δp38 mutant had a modest replication defect (FIG 15C), but the Δp36 mutant and the SynI1 mutant (that lacks p33) replicated with wt kinetics (FIG 15B and 19A). These data suggested that p33 and p36 are completely dispensable for replication in this cell type. To test the importance of the E1 p33, p36, and p38 isoforms in other cell types, viral replication kinetics in two additional cell types permissive for MCMV replication was analyzed. In SVEC4-10 endothelial cells, the mutant MCMVs replicated with similar kinetics as in NIH-3T3 fibroblasts. The Δp36 and SynI1 mutants attained titers comparable to the parental wt virus, but the Δp38 mutants had a modest replication defect (FIG 20A). In RAW267.4 macrophages only MCMV Δp36

replicated to wt titers (FIG 20B) while MCMV Δp38 displayed a pronounced replication defect. Surprisingly, a similar replication defect was observed for the MCMV SynI1 mutant, suggesting an important role for either the E1 p33 isoform or the authentic intron 1 for MCMV replication in macrophages. In lysates of infected SVEC4-10 and RAW267.4 cells expression of the expected E1 protein isoforms was detected by immunoblot analysis (FIG 20C and D). Interestingly, an additional band was detected above the band representing the p33 isoform in SVEC4-10 endothelial cells (FIG 20C). The presence of this band correlated with the presence of the p33 isoform, suggesting that a post-translational modification of p33 occurring in SVEC4-10 but not in NIH-3T3 cells might be responsible. A similar protein band was detected by the E1 antiserum in RAW267.4 macrophages. However, the same band was present in mock-infected macrophages, indicating that in those cells this band does not represent an MCMV protein.

From those data it can be concluded that the E1 p38 isoform is important for efficient replication in all cell types tested, whereas the E1 p36 appears to be dispensable in these cell types - at least *in vitro*. Although dispensable for MCMV replication in fibroblasts and endothelial cells, the E1 p33 isoform or the authentic M112/113 intron 1 are required for efficient replication in RAW264.7 macrophages.

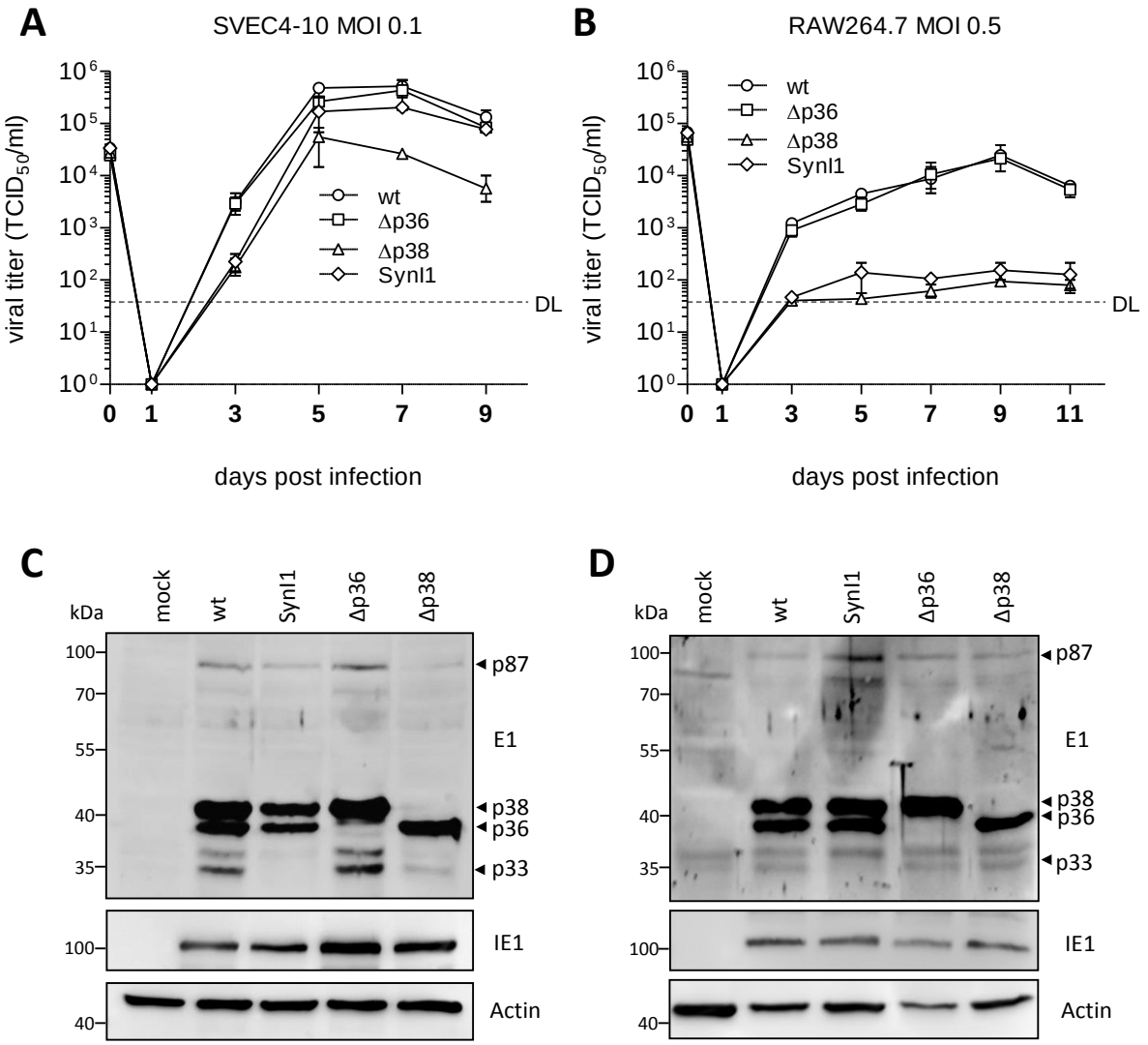


FIG 20. Replication properties of MCMV M112/113 mutants in endothelial cells and macrophages. (A) SVEC4-10 endothelial cells were infected with an MOI of 0.1 $TCID_{50}$/cell, (B) RAW264.7 macrophages were infected with an MOI of 0.5 $TCID_{50}$/cell with MCMV mutants. Viral replication kinetics were determined as described above. (C) SVEC4-10 cells or (D) RAW264.7 were infected at an MOI of 3 $TCID_{50}$/cell with the indicated viruses. E1 protein expression was determined by immunoblotting. IE1 was used as infection control and β-actin was used as a loading control.

5.7 Affinity purification and mass spectrometry analysis identify potential interaction partners of E1p87

The influence of the different isoform on MCMV replication could be identified using virus mutants having individual E1 isoforms inactivated. However the exact function(s) of those isoforms remains to be elucidated. A first step towards deciphering the functions of the E1 proteins is to find cellular or viral proteins interacting with them in the context of a MCMV infection. IP of the E1 proteins of SILAC NIH-3T3 infected with MCMV combined with mass spectrometry was applied to find potential interaction partners. To this end, MCMV mutants with N-terminally and C-terminally HA tagged E1 proteins were constructed. Between the E1 coding sequence and the HA tag a tobacco etch virus protease cleavage site (tev) was introduced. The HA tag allowed high yield affinity purification whereas the tev cleavage site can be used to cut off the tag from the protein in order to specifically elute the target protein from the affinity support (FIG 21A).

Surprisingly, the N-terminally tagged MCMV mutants could not be reconstituted in numerous attempts. This suggests an essential role for the N-terminus of one or more E1 isoforms either for protein function or correct folding. In contrast the C-terminally tagged HA MCMV E1 mutant was easily reconstituted and grew to wt titers. This MCMV mutant was used to infect heavy-labeled NIH-3T3 cells. Light-labeled NIH-3T3 cells were infected with wt MCMV as control. Potential interaction partners were identified after affinity purification of the large E1p87 isoform (FIG 21B) using MS analysis (done by Stefan Loroch, ISAS in Dortmund, Germany).

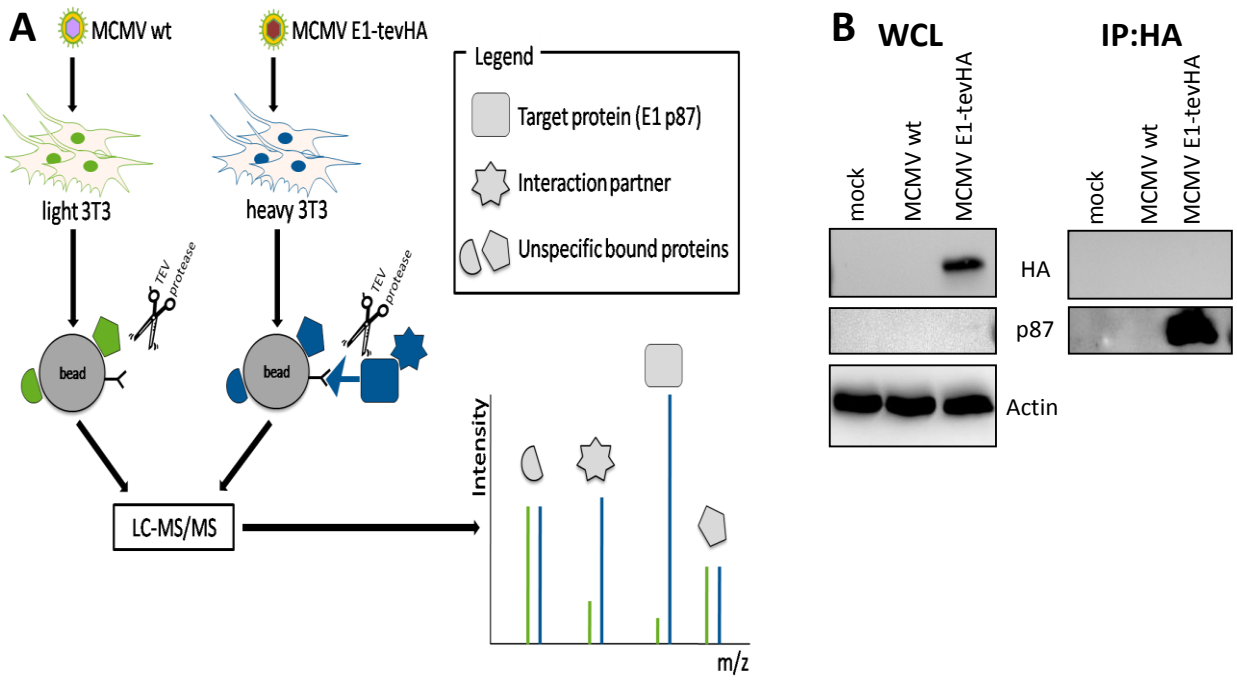


FIG 21. Affinity purification of MCMV E1p87. (A) Schematic showing the strategy to find interacting proteins of the large E1p87 isoform in infected SILAC NIH-3T3 cells. Modified from https://www.broadinstitute.org. (B) Immunoblot showing the whole cell lysates (WCL) and affinity purified E1p87 after AcTEV protease elution (IP:HA). NIH-3T3 cells were infected using the indicated viruses at an MOI of 5 $TCID_{50}$/ml. Blots were stained using anti HA, anti E1 and anti-actin primary antibodies. Actin was used as loading control.

The interaction partners were ranked according to the number of unique peptides identified and the normalized ratio of heavy over light counts (Table 1). Interestingly, most of the identified potential interactions partners were from the group of heterogeneous ribonucleoproteins (hnRNPs). The three highest ranking potential interaction partners, namely RNA Binding Motif Protein, X-Linked (RBMX or hnRNP G), hnRNP A1 and hnRNP A3 were further investigated

Table 1 Potential interaction partners of MCMV E1p87. The table shows the number of unique peptides, peptide spectrum match (PSMs) and the enrichment compared to control of potential interaction partners of E1p87.

Description	# Unique Peptides	# PSMs	Enrichment
RBMX (hnRNP G)	4	6	8,136
Putative RNA-binding protein 3	1	1	5,289
Nuclear fragile X mental retardation-interacting protein 2	1	1	3,546
Ataxin-2-like protein	1	1	2,569
Heterogeneous nuclear ribonucleoprotein A3	4	4	2,321
Heterogeneous nuclear ribonucleoprotein A1	4	7	1,855
Zinc finger CCCH-type antiviral protein 1	1	1	1,773
Retinoic acid-induced protein 1	1	1	1,703
Heterogeneous nuclear ribonucleoprotein M	9	13	1,700
Heterogeneous nuclear ribonucleoproteins A2/B1	3	4	1,655
Signal recognition particle 68 kDa protein	1	1	1,654

5.8 Co-IP cannot conclusively confirm hnRNPs as interaction partners of E1p87

In order to confirm the potential interaction of E1p87 with hnRNP A1, hnRNP A3 and/or RBMX, NIH-3T3 cell lines stably expressing flag-tagged versions of those proteins were engineered using retroviral transduction and puromycin selection. This strategy was chosen, because all tested commercially available antibodies against any of those proteins failed to reproducibly detect them in WB experiments. After transduction and selection, all cell lines showed stable expression of the respective flag-tagged proteins.

If E1p87-HA was affinity purified no interaction with flag-tagged hnRNP A1 or hnRNP A3 could be detected (FIG 22A). Also if flag tagged hnRNP A1 or hnRNP A3 were affinity purified no E1p87-HA co-precipitated (FIG 22B). An interaction with hnRNP A1 and/or A3 with E1p87 could not be confirmed. In contrast if E1p87-HA was affinity purified from infected flag-tagged RBMX expressing cells, a specific signal for flag-RBMX could be detected after long exposure times (FIG 22A). Using an MCMV mutant carrying an HA-tagged version of M45 it could also be shown that the observed interaction of E1p87 and RBMX does not appear using any viral HA-tagged protein since RBMX does not associate with tagged M45 (FIG 22C). Consequently RBMX does not simply co-precipitate with any HA-tagged protein. However when the experiment was done vice versa, meaning flag-tagged RBMX was affinity purified, the interaction between E1p87 and RBMX could not be confirmed.

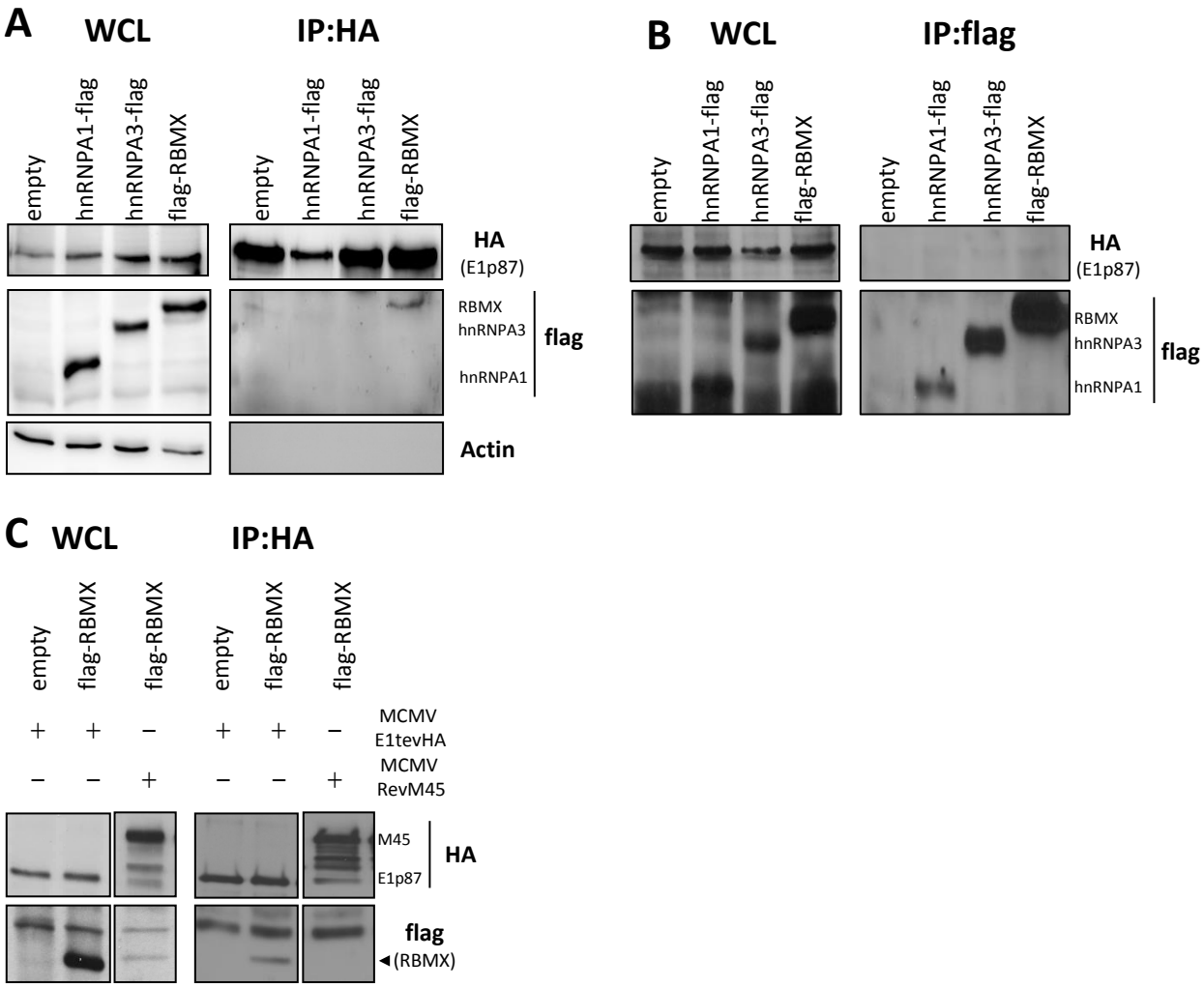


FIG 22. CoIP of E1p87 and potential interacting hnRNPs. Immunoblot showing the whole cell lysates (WCL) and affinity purified (A) E1p87tevHA (IP:HA) or (B) flag-tagged hnRNPs (IP:flag). NIH-3T3 cells stably expressing the indicated hnRNPs were infected using MCMV E1tevHA at an MOI of 5 $TCID_{50}$/ml. (C) NIH-3T3 stably expressing RBMX or NIH-3T3 transduced with empty vector were infected using either MCMV E1tevHA or as control MCMV RevM45 at an MOI of 5 $TCID_{50}$/ml. Blots were stained using anti-HA, anti-flag and anti-actin primary antibodies. Actin was used as loading control in one experiment.

Taken together these results strongly argue that these tested potential interaction partners of E1p87 are false positives. This notion is further underlined by the fact that the human homologs of those 3 proteins appear frequently in affinity purification experiments [132]. RBMX was detected in 169 out of 411 Co-IP studies, whereas hnRNP A1 and hnRNP A3 were detected in 268 and 183 out of those 411 studies, respectively.

5.9 HCMV E1 isoform inactivation shows similar impact on viral replication compared to MCMV E1 isoform inactivation.

To compare the MCMV M112/113 gene products with its HCMV counterparts, similar mutations as described in section 5.2 and 5.5 were introduced in UL112/113. This was done for both overexpression plasmids as well as HCMV mutants.

All HCMV E1 mutants were based on an HCMV TB40/E carrying N-terminally HA tagged E1 proteins made by Rebekka Brost. Interestingly an N-terminal HA tag does not interfere with viral replication in HCMV. This is surprising hence in MCMV, N-terminally tagging the M112/113 gene products completely abolishes viral replication. In contrast a C-terminal HA tag does not interfere with viral replication in both MCMV and HCMV.

To inactivate the individual E1 isoforms mutations were introduced as follows (FIG 23): Intron 1 was removed in the Δp34 mutant, leading to a fusion of exons 1 and 2. In the Δp43 mutant the splice acceptor A2 was mutated by one nucleotide exchange that does not affect the amino acid sequence encoded by the unspliced transcript. In the same manner, splice acceptor A3 was mutated to generate the Δp50 mutant. To eliminate the expression of the large p84 isoform, one point mutation was introduced shortly after splice donor D2. This point mutation leads to a premature stop codon, which should result in a HCMV mutant containing a truncated p84 protein lacking the last 336 amino acids.

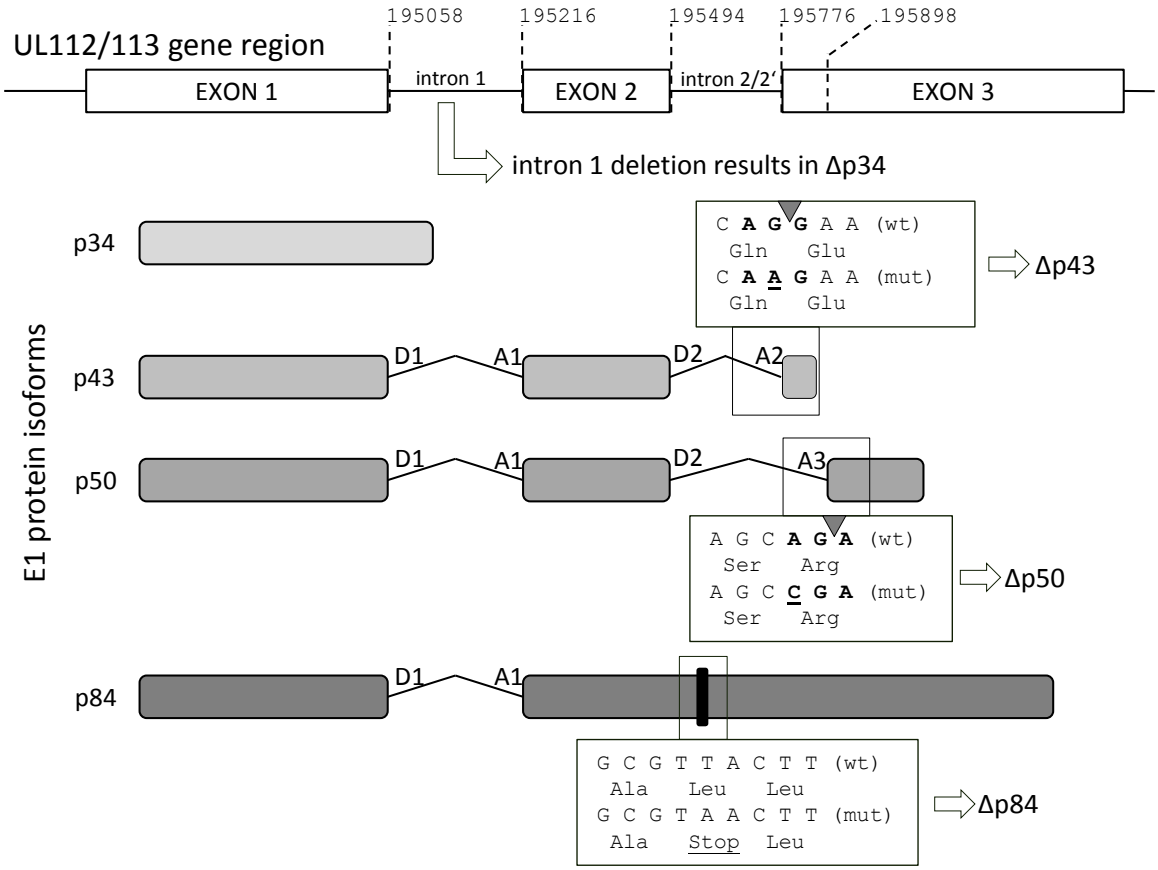


FIG 23. Mutagenesis of the UL112/113 locus. Schematic view of the HCMV UL112/113 locus and the four proteins expressed from this locus by alternative splicing. Mutations introduced to inactivate the expression of individual protein isoforms are shown. Splice donor and acceptor sites are indicated by arrow heads, and the minimal consensus sequences are shown in bold. Nucleotide and amino acid changes are underlined.

To determine whether the mutations lead to a loss of individual E1 protein isoforms, the mutated plasmids were transfected into 293A cells and their cell lysates were analyzed by WB using an anti-HA antibody. Plasmids containing individual HCMV E1 isoform have been used as control. Despite an overall weak expression of E1p34 and especially E1p50, all plasmids showed inactivation of the desired E1 isoforms (FIG 24).

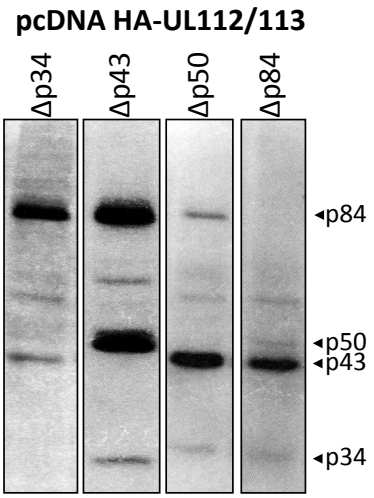


FIG 24. HCMV E1 isoform expression from E1 inactivation plasmids. 293A cells were transfected with empty vector or expression plasmids containing HA-tagged mutant UL112/113 sequence. E1 protein expression was detected by immunoblotting using an HA-specific antibody. This figure shows sections of blots containing multiple clones of one plasmid.

The HCMV E1 inactivation mutant BACs plus a HCMV E1 stop mutant with 3 premature stop codons at amino acid position 9, 11 and 12 were subsequently transfected in MRC-5 cells. Following this, the cells were checked for GFP positive foci as well as cytopathic effects (CPEs) which are an indicator of HCMV reconstitution from BAC. The time of occurrence and speed of spreading of those CPEs was compared to HCMV TB40/E wt BAC transfected cells to estimate the viral replication kinetics of the mutants.

The HCMV E1 stop mutant could not be reconstituted after transfection. However, the HCMV Δp50 mutant could be easily reconstituted and showed fast growth and big plaque size. Both HCMV Δp84 and HCMV Δp43 reconstitution from BAC transfected cells proved to be very difficult and both virus mutants showed strongly impaired replication. In fact HCMV Δp43 replicated so slowly that a sufficient amount of infected cells for immunoblotting could not be obtained. The HCMV Δp34 mutant could not be reconstituted at all, but this phenotype was rescued if the splicing of UL112/113 intron 1 was restored by an artificial intron using the same strategy as described for MCMV in section 5.5. Interestingly the p34 isoform shows a stronger expression when expressed during HCMV infection compared to transfected cells (FIG 25), whereas the overall weak expression of p50 and p84 remains. In general the different expression of the HCMV E1 isoforms in infected cells confirms published data [67, 72].

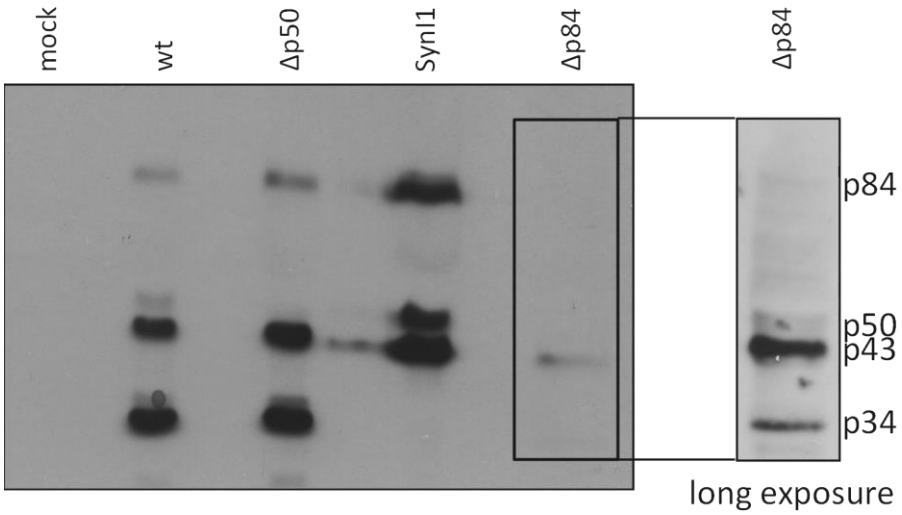


FIG 25. Protein expression of HCMV UL112/113 mutants. MRC-5 cells were infected at an MOI of 3 $TCID_{50}$/cell with HCMVs containing HA-tagged wt or mutant UL112/113. E1 isoform expression was determined using an anti-HA antibody.

In summary, these data show strong parallels of the importance of the different HCMV E1 isoforms compared to their MCMV counterparts (Table 2). However more experiments need to be conducted to quantify the observed growth defects. Constructing HCMV E1 mutants being inactivated for two or more E1 isoforms simultaneously as well as HCMV E1 revertants are needed for further studies of the HCMV E1 isoforms. In addition affinity purification to find interaction partners of HCMV E1 isoforms need to be done in future to describe the impact of the individual HCMV E1 isoforms in more detail.

Table 2. Comparison of HCMV and MCMV mutants inactivated for individual E1 isoforms.

HCMV mutant	replication properties	MCMV homolog	replication properties
E1 stop	no replication	E1 stop	no replication
E1tevHA	like wt	E1tevHA	like wt
HAtevE1	like wt	HAtevE1	no replication
Δp34	no replication	Δp33	no replication
Δp50	like wt	Δp36	like wt
Δp43	severely impaired	Δp38	impaired
Δp84	severely impaired	Δp87	no replication
SynI1	like wt	SynI1	like wt

6 Discussion

6.1 Revision of the MCMV M112/113 gene region splicing pattern

Using RNA-Seq to identify splice junctions in MCMV transcripts, the M112/113 splicing pattern was revised. Surprisingly the previously published splicing pattern of this region was not fully correct [66, 68]. Instead of sharing a common splice donor the two medium sized isoforms E1p36 and E1p38 share a common splice acceptor and make use of two different donor sites (FIG 9). Interestingly this distinguishes the MCMV M112/113 splicing pattern from the UL112/113 splicing pattern [67, 133] (FIG 23). These findings underline that even if the overall structure of a CMV gene region looks the same there can be subtle differences between CMVs from different species. Consequently, often the information gained from investigating CMV of one species cannot simply be transferred to CMVs of others species without additional testing.

The exact genomic positions of the M112/113 splice sites were revised using RNA-Seq and confirmed by mutagenesis of M112/113 plasmids (FIG 11 and 12). While completely abrogating expression of the intentionally inactivated E1 isoform(s), the introduced changes did not drastically change expression of the other respective isoforms. In general, the small E1p33 and the large E1p87 isoforms were expressed to lower levels than the two medium sized E1 isoforms. Strikingly this is also reflected in the abundance of spliced compared to unspliced M112/113 transcripts in infection (FIG 10). This shows that splicing of the M112/113 transcripts is a very efficient mechanism and the retention of introns is a rare event, consequently leading to more protein from fully spliced transcripts compared to protein translated from partially spliced or non-spliced transcripts. Remarkably this does not seem to be the case for the UL112/113 gene products. The HCMV E1p50 is expressed very weakly despite the fact that it is fully spliced, whereas the E1p34 is strongly expressed even though it is derived from a completely unspliced transcript. Whether this is due to the involvement of additional regulating factors in the case of HCMV, a difference in the protein turnover rates of the HCMV and MCMV E1 isoforms or rather a general difference between HCMV's and MCMV's spliced gene products remains to be elucidated.

Besides the bands for the canonical E1 isoform several other weak bands could be detected in both transfection and infection experiments. The two most likely explanations for those bands are that they either represent degradation products of the

E1 isoforms or modified versions of the canonical isoforms, e.g. due to phosphorylation as it has been reported for HCMV UL112/113 or other post translational modifications [133]. In fact, two recent publications reported the detection of 3 small HCMV UL112/113 proteins, which were apparently generated by calpain-mediated cleavage [134]

Nevertheless it cannot be fully ruled out that (some of) those bands represent other E1 isoforms arisen from transcripts using a different splice pattern than the canonical ones. However this hypothesis is not supported by RNA-Seq data, since other unknown splice junctions were detected at a very low frequency, comprising less than 1 % of the splices in this region. Most of these rare splices were not conserved between the MCMV strains Smith and K181. However, one putative unknown intron of 214 nucleotides for MCMV Smith and 217 nucleotides for MCMV K181 was detected 180 bp downstream of the A2 acceptor site. Though, the used splice donor and acceptor sites do not match the minimal consensus sequences in both strains. Hence it remains unclear whether this is a true non-canonical intron used at a low frequency or rather an artifact. The latter is being supported by the fact that the predicted size of this intron differs between the two MCMV strains, which was not observed for the introns of the confirmed isoforms. In an RNA-seq analysis of the HCMV genome such atypical splice junctions were classified as probable artifacts [105].

6.2 Inactivation of MCMV E1 isoforms in infection

With the help of the MCMV E1 stop mutant it could be shown that the M112/113 gene products are essential for viral replication confirming previously published data [76]. Astonishingly N-terminal but not C-terminal tagging of the E1 isoforms abrogates MCMV replication. Whether the N-terminal HA tag interferes with a critical domain at the N-terminus of the MCMV E1 proteins or the tag disturbs the fold of the MCMV E1 isoforms remains speculative and the objective of future experiments. The interaction of the UL112/113 gene products among each other as well as its interaction with IE2 has been mapped to the N-terminal part of the HCMV E1 isoforms [82]. However, HCMV mutants tolerate both N-terminally and C-terminally tagged E1 proteins.

BAC mutagenesis also enabled me to inactivate individual E1 isoforms in MCMV infection. Besides proving the principle that splice site mutagenesis of an alternatively spliced gene is a valuable tool to elucidate the influence of particular isoforms, it also allowed to further describe the impact of individual MCMV E1 isoforms on viral

replication. The following paragraphs discuss the consequences of the inactivation of individual E1 isoforms in more detail.

6.3 Inactivation of MCMV E1p33 and deletion of M112/113 intron 1

Surprisingly, the deletion of M112/113 intron 1, which eliminates the small E1p33 isoform, was not compatible with MCMV replication. The E1p33 protein is translated from transcripts retaining intron 1 and differs only by a single C-terminal lysine residue from the N-terminal parts of E1p36, E1p38, and E1p87. Viral replication was not rescued by ectopic expression of E1p33, but by insertion of a synthetic intron in place of intron 1. The MCMV SynI1 mutant replicated to wt levels in NIH-3T3 fibroblasts and SVEC4-10 endothelial cells (FIG 19A and 20A). MCMV SynI1 infected cells did not show expression of E1p33 (FIG 19B and 20C), nor were unspliced mRNA transcripts detected (FIG 19C). Therefore, it is the presence of intron 1 rather than the expression of E1p33 that is essential for MCMV replication.

Currently, it is unknown which essential function intron 1 fulfills. One possible function of intron 1 could be to ensure sufficient expression of the essential large E1p87 isoform. In the absence of intron 1, the E1p87 protein is encoded by an unspliced transcript. Unspliced transcripts are exported with low efficiency from the nucleus by the cellular mRNA export machinery [135, 136]. Therefore herpesviruses encode for export factors ensuring the export of the mainly unspliced viral transcripts [137-140]. For HCMV it was shown that UL69 acts as such an export factor and effectively mediates the export of unspliced mRNAs from the nucleus to the cytoplasm [141, 142]. Even if such an export factor has not yet been described [143] one can assume that a similar machinery exists in MCMV. However for HCMV UL69 it was shown that it is expressed with early-late kinetics and therefore after the expression of the UL112/113 gene products. Consequently a similar export factor in MCMV might not be present when the E1p87 transcript needs to be exported from the nucleus and export of those transcripts relies on the cellular export machinery. Accordingly, removing intron 1 from the E1p87 transcript might result in a very inefficient export of this transcript resulting in a loss of the encoded essential protein. The fact that an E1 expression plasmid without intron 1 expresses robust amounts of p87 (FIG 12) does not refute this hypothesis. Expression plasmids with very strong promoters lead to high-level protein expression even from unspliced transcripts, and this might mask a reduced export of unspliced transcripts. Another possible explanation for the requirement of intron 1 is an effect on splicing of

the second intron, leading to an altered ratio of the p36, p38, and p87 isoforms. Indeed, in cells transfected with the M112/133 Δp33 plasmid, expression of p38 was strongly reduced (FIG 12).

The inability of MCMV SynI1 to efficiently replicate in RAW264.7 macrophages suggests an additional or altered function of the E1p33 protein in these cells. This function, making E1p33 an interesting candidate for a cell tropism factor, remains to be elucidated in future experiments.

6.4 Inactivation of MCMV E1p36 and E1p38

The inactivation of the E1p36 isoform seems to be dispensable for viral replication in all tested cell lines (FIG 14B, 20A and B), whereas the inactivation of E1p38 results in replication deficient MCMV mutant in all cell lines tested. Strikingly the simultaneous knockout of both medium sized isoforms results in an MCMV mutant unable to replicate. This indicates that the MCMV medium sized E1 isoforms have similar functions in virus biology with E1p38 being more and E1p36 being less efficient. This is not surprising taking into account that both medium sized isoforms show 98.7 % of sequence identity. Nevertheless E1p36 seems to be able to partly compensate the loss of E1p38. By chance, a mix of two MCMV quasispecies was discovered; both being inactivated for the E1 medium sized isoforms. Interestingly one of the MCMV mutants expresses a truncated E1p87 isoform which compensates the loss of the two medium sized E1 isoforms. This unconventional medium sized E1 isoform is result of a unavoidable premature stop codon leading to the loss of the full length E1p87 isoform. The other MCMV mutant in this mix of viruses provides the essential E1p87 isoform. The fact that these two mutants could neither be separated nor reconstituted from BACs individually supports the concept (FIG 17).

This allows speculating that any medium sized E1 isoform having sequence identity with the N-terminal part of either E1p36/E1p38 is sufficient to allow MCMV replication as long as E1p87 and the M112/113 intron 1 are present. Additionally this finding nicely shows the capability of two distinct virus mutants to compensate each other's replication defects when infecting the same cells, which might display a general feature of virus evolution as it has been described for RNA viruses [144].

In general the introduction of nucleotide exchanges to inactivate a splice donor or acceptor sites is probably the smallest possible alteration one can introduce. Nevertheless, such a small mutation may still have an impact on the expression of

proteins derived from the remaining splice variants. The inactivation of the p33, p36, or p38 isoform did not result in strong changes in the abundance of the remaining protein isoforms (FIG 14D). However, inactivation of the two most abundant isoforms, p36 and p38, by mutation of the shared splice acceptor site did result in a strong increase of the p33 and p87 isoforms in the spontaneous mutant (FIG 16B), which are usually expressed only at low levels. The increase in expression of those two isoforms is most likely due to an increase of the remaining two transcripts, assuming an overall constant number of total transcripts from the M112/113 locus. This observed shift in abundance was not unexpected as a previous study has found that elimination of MCMV IE1 transcript by fusion of exons 3 and 5 results in an enhanced expression of IE3 [145].

6.5 Inactivation of MCMV E1p87

Inactivation of the E1p87 isoform proved to be most difficult. The only mutation that prevents E1p87 expression but leaves the coding sequence of the other isoforms intact is a nonsense mutation within intron 2 (FIG 11). This mutation should lead to the expression of a substantially truncated E1p87 protein; however such a protein was not identified in cells transfected with a mutant M112/113 expression plasmid (FIG 12). The truncated protein might be unstable or masked by the E1p36 or E1p38 proteins, which are similar in size. The most obvious explanation for the inability of the MCMV Δp87 to replicate is an essential function of E1p87.

As a matter of fact, C-terminal truncation of the UL112/113 p84 isoform, the homolog of p87 in HCMV, also resulted in a non-replicating virus [85]. This replication defect is probably due to the inability of the truncated p84 protein to recruit the DNA polymerase accessory protein, pUL44, to viral pre-replication compartments [85]. Although such a function has not yet been described for MCMV p87, this protein was reported to control the repressive effect of IE3 on the MIE promotor [87], and there might be additional functions exclusive to E1p87. Unfortunately, *in silico* protein domain predictions did not yield any informative results for the unique C-terminal part of E1p87. In any case, the possibility that the truncated E1p87 protein prevents MCMV replication by a dominant negative effect cannot be ruled out entirely. As the truncated E1p87 protein differs from E1p38 only by 12 amino acids, this possibility appears rather remote. However, an attempt to rescue the loss of E1p87 by expression of the protein in *trans* was not successful even though E1p87 could be detected by immunoblot in stably transduced cells (FIG 15). Whether the retroviral transduced E1p87 ORF did not lead to the

expression of sufficient amounts of the protein, the expression of E1p87 in those cells was not stable enough or had a wrong temporal regulation, remains to be investigated.

The importance of the large E1p87 isoform for viral replication was additionally underlined by the emergence of two MCMV E1 mutants, which complemented each other and have been discussed in a previous paragraph (FIG 16 and 17). One mutant virus expressed a truncated version of E1p87 apparently compensating for the loss of the two medium-sized isoforms, whereas the other mutant expressed full-length E1p87. The fact that these two mutants could not be separated from each other supports the concept that the full-length E1p87 as well as one medium sized isoform are required for MCMV replication.

6.6 Potential interaction partners of E1p87

A way to investigate the function of a viral gene product is to find its cellular interaction partners or viral interaction partners with known functions. This has been tried for MCMV E1p87 using SILAC NIH-3T3, affinity purification and LC-MS/MS (FIG 21). Similar strategies proved to be successful in the past [130, 146].

Strikingly no homologs of viral proteins described to interact with HCMV E1p84 such as M44 (homolog of UL44 [85, 147]) were detected in the screen. A protein-protein interaction between the E1 proteins and IE3 as it has been described by Martinez and colleagues could also not be confirmed [148]. Moreover the screen does not indicate interactions among E1p87 and the other E1 isoforms as described for HCMV E1p84 [84].

Besides lacking any potential viral interaction partners, the SILAC screen revealed three potential cellular interaction partners of E1p87: hnRNP A1, hnRNP A3 and RBMX (hnRNP G) (Table 1). Interestingly all of them belong to the family of heterogeneous nuclear ribonucleoproteins (hnRNPs) and are described to be present in the nucleus of human cells [149-151]. This would allow them to interact with E1p87 which has also been reported to be localized in the nucleus [66].

The family of hnRNP proteins is very diverse and most of their members are associated with the processing and regulation of mRNAs [152]. Strikingly other functions unrelated to the mRNA regulation have been described for all three of them. RBMX has been reported to be involved in the regulation of apoptosis as well as DNA damage response [151, 153]. Despite being crucial for pre-mRNA splicing, the multifunctional hnRNP A1 is involved in processing of microRNAs, telomere maintenance and the regulation of transcription factor activity [154]. Also for hnRNP A3 which is by far the least studied of

those three proteins, a function in maintenance of telomeric repeats has been described [155]. Among those very diverse functions several are also important for CMV replication, therefore an interaction of E1p87 with one or more of these cellular proteins might influence their functions and by that facilitate CMV replication.

However, additional IP experiments using NIH-3T3 with tagged versions of these proteins could not convincingly verify any of the three proteins as an interaction partner of E1p87 (FIG 22). For hnRNP A1 and hnRNP A3 no interaction with E1p87 was shown. The potential RBMX E1p87 interaction could only be shown when E1p87 was affinity purified using very long exposure times for the immunoblot. Conversely RBMX E1p87 interaction could not be detected after affinity purification of RBMX. While E1p87 does not seem to interact with hnRNP A1 and A3 at all, the interaction with RBMX seems to be very weak. Taken together with the fact that all three of these proteins are frequently detected as false positive in different affinity purification experiments, it may be concluded that the presumed interaction of E1p87 with these three interaction partners is a false positive or an experimental artefact [132].

As the presented approach was unsuccessful, no additional function of E1p87, besides relieving the repressive effect of IE3 on its own promotor, could be described [87]. In the future alternative approaches are needed to decipher those functions. Chemical crosslinking of E1p87 and its potential interaction partners may be an option; by using this method one could overcome the technical problem that treatment with harsh buffers that are able to lyse the nucleus might destroy protein-protein interactions. Another option would be the recently developed BioID method [156]. This method utilizes the transfer of biotin to proteins in close proximity of the target protein. These biotinylated proteins can then be identified by MS. Besides abolishing any lysis buffer related destruction of interactions this method can also be used to detect transient interactions among proteins. Since it has been reported that the E1 proteins can bind single and double stranded nucleic acids, chromatin IP experiments addressing those feature might also be of future interest [69, 70, 79].

6.7 Comparison of the inactivation of HCMV UL112/113 gene products with the inactivation of MCMV M112/113 gene products

CMVs can only infect cells of their natural host or very closely related species. Therefore MCMV is used as an *in vivo* model for HCMV infection [30]. Both viruses are very similar in terms of their biology and many of their homologous gene products do have similar or

even identical functions for example the MCMV M50 and HCMV UL50 [146]. Nevertheless it is crucial to investigate whether a feature described for MCMV can also be found in HCMV. Therefore splice site mutagenesis of the HCMV UL112/113 gene region was performed (FIG 23, 24 and 25) and the resulting data were compared to those acquired using MCMV M112/113 splice site mutants (Table 2). Since experiments using UL112/113 splice mutants are currently still ongoing, conclusions drawn from the results obtained so far must be considered preliminary. However, even the first set of results allows comparison of the importance of HCMV UL112/113 and MCMV M112/113 proteins for viral replication.

Both HCMV and MCMV E1 stop mutations result in viruses unable to replicate, suggesting an essential role for the viral E1 proteins in both viruses. Surprisingly, studies by Dunn et al. and Yu et al. suggest a rather augmenting role of the UL112/113 gene region for HCMV replication [74, 75]. If UL112 or UL113 were knocked out in those studies HCMV replication was severely impaired but detectable. In contrast to my work those studies used the laboratory adapted strains Towne and AD169 whereas I made use of HCMV strain TB40/E, which is closer to clinical CMV isolates [157]. It remains to be elucidated whether the UL112/113 gene region is uniquely essential in TB40/E or whether there is a general strain specific difference of its impact on HCMV replication.

As for MCMV *en passant* BAC mutagenesis allowed inactivation of individual HCMV E1 isoforms to study their impact on viral replication. Using these virus mutants, the HCMV E1p50 isoform could be described as dispensable for HCMV replication similar to MCMV E1p36. In contrast, inactivation of HCMV E1p43 resulted in a growth defect similar to the inactivation of MCMV E1p38. This E1p43 related phenotype may be caused by reduced transactivation of the UL54 promotor hence E1p43 is associated with transcriptional regulation of UL54 both dependent and independent of IE2 [81, 126].

Interestingly the inactivation of HCMV E1p84 resulted in severely impaired HCMV replication. This is in contrast to MCMV where the inactivation of E1p87 resulted in a virus incapable of replication, indicating that the function(s) of the large E1 isoform are of different importance in HCMV and MCMV. One can speculate that the loss of E1p84 can be partly compensated by E1p43 and/or E1p50, restoring for example the recruitment of UL44 to viral pre-replication foci [82, 85]. In MCMV however, the two medium sized isoforms may not be able to compensate E1p87 functions such as regulating IE3 transcription [87]. Simultaneous inactivation of HCMV E1p84 together

with E1p43 and/or E1p50 can help to answer the question if one or both of the medium sized isoforms can (partially) take over E1p84's function(s). In contrast to the rather augmenting nature of E1p84 suggested by the results I obtained, it has been described in another study that a truncated E1p84 in HCMV strain AD169 results in a replication defective virus [85]. Whether this is a strain specific effect or due to other not detected compensatory mutations needs to be investigated in future experiments.

Intriguingly intron 1 of both UL112/113 and M112/113 seems to be indispensable for CMV replication, whereas the expression of the smallest E1 isoform is not needed for HCMV and MCMV replication in fibroblasts. It is very likely that for HCMV this is due to similar reasons as for MCMV, as described previously. It is remarkable that splicing of a certain gene region but not its gene product appears to be essential for conserved viral replication among at least two members of the CMVs. This could be a hint that CMVs or even herpesviruses in general use splicing to ensure the expression of critical proteins at times when viral mRNA export factors are not yet present. This hypothesis would also explain why many of the spliced transcripts are expressed during latency or at the very beginning of the lytic cycle. It could also be possible that herpesviruses use (alternative) splicing as a means to regulate the levels or activities of viral proteins.

6.8 Concluding remarks

In summary the data in this thesis demonstrate that single viral protein isoforms can be inactivated individually without affecting other gene products derived from the same gene region. This can be achieved by applying splice donor or acceptor site mutations, intron deletion, and substitution mutagenesis. Although mutagenesis of a splice donor or acceptor site has been done before on a few occasions, e.g. for the inactivation of the stable introns derived from the HSV-1 latency associated transcript or CMV immediate-early transcripts [158-160], to the best of knowledge this is the first comprehensive mutational analysis of a differentially spliced herpesvirus gene. Moreover the methods presented here for splice site mutagenesis are not just limited to cytomegaloviruses or herpesviruses and can in principle be transferred to other DNA viruses, making them a valuable tool for virus research in general.

Using these techniques the importance of individual isoforms of the MCMV M112/113 region as well as the HCMV UL112/113 region for viral replication were investigated. Splice site mutagenesis helped describe similarities and differences in the importance of

single E1 isoforms for those two viruses. The data clearly shows that the four different E1 isoforms are not functionally redundant neither in MCMV nor in HCMV. Inactivation of a single E1 isoform can have different consequences for viral replication, with some of the E1 proteins being dispensable, others augmenting or even essential for viral replication. Furthermore, I could convincingly demonstrate that the splicing of the M112/113 and UL112/113 intron 1 itself, but not its coded protein product, is essential for replication of both MCMV and HCMV. This substantiates that splicing is a regulatory mechanism used by the virus to control the expression of viral proteins at times when no other viral regulatory factors are present.

Even though the SILAC based affinity purification of the large MCMV E1p87 isoform did not result in the identification of cellular interaction partners, improvements of this method may help to find interaction partners of E1p87 in the future. This would be an important step towards a better understanding of the function(s) of E1p87, the most important of the E1 proteins.

To summarize, this study provides more detailed understanding of the role of E1 proteins in CMV replication. For the first time, all the different E1 isoforms have been investigated systematically to decipher their importance in CMV replication. Thereby this study provides novel insights into the role of this important set of viral proteins and is a major step forward in understanding the function(s) of these surprisingly less studied viral gene products.

7 Material

7.1 Cell lines

name	description	reference / source
10.1	Murine embryonal fibroblasts from BALB/c, spontaneous immortalized, defect in p53	[161]
HEK-293A	Human embryonal kidney epithelial cells, subclone of the 293 cell line, selected on flat morphology	Invitrogen (R705-07)
MRC-5	Primary human lung fibroblasts, only limited passaging possible	ATCC (CCL-171)
NIH-3T3	Murine embryonal fibroblasts from NIH/Swiss, spontaneous immortalized	ATCC (CRL-1658)
NIH-3T3 flag-RBMX	NIH-3T3 stably transduced with flag-RBMX expressing pMSCVpuro	this study
NIH-3T3 hnRNPA1-flag	NIH-3T3 stably transduced with hnRNPA1-flag expressing pMSCVpuro	this study
NIH-3T3 hnRNPA3-flag	NIH-3T3 stably transduced with hnRNPA3-flag expressing pMSCVpuro	this study
Phoenix	Retroviral packaging cell line stably expressing gag, pol and env; based on HEK-293T cells	[162]
RAW264.7	Murine macrophage-like cells, Abelson murine leukemia virus transformed	ATCC (TIB-71)
SVEC4-10	Murine endothelial cells, immortalized using SV40 large T antigen	ATCC (CRL-2181)

7.2 Viruses

name	description	reference / source
HCMV HA-tev-E1	HCMV TB40/E GFP with N-terminal HA-tev tagged E1 proteins	Rebekka Brost
HCMV TB40/E GFP	HCMV TB40/E strain, cloned as BAC, expressing GFP	[163]
MCMV K181	MCMV K181 strain, cloned as BAC	[164]
MCMV RevM45	MCMV Smith pSM3fr-MCK-2fl with reinserted M45 full length ORF including an C-terminal HA tag	[165]
MCMV Smith GFP	MCMV Smith strain, cloned as BAC, expressing GFP	[166]

The following viruses were generated during the course of this work

name	description
HCMV E1 Stop	HCMV HA-E1 with 2 point mutations introduced to create 2 premature stops in UL112/113, Y11Stop, W12Stop
HCMV SynI1	HCMV HA-E1 having the UL112/113 intron 1 replaced by a synthetic intron (SynI)
HCMV Δp34	HCMV HA-E1 with deletion of the UL112/113 intron 1, fusing UL112/113 exon 1 and 2
HCMV Δp43	HCMV HA-E1 with inactivated E1p43 by mutation of the splice acceptor A2
HCMV Δp43Δp50	HCMV HA-E1 with inactivated E1p50 and E1p43 by mutation of the shared splice donor D3
HCMV Δp50	HCMV HA-E1 with inactivated E1p50 by mutation of the splice acceptor A3
HCMV Δp84	HCMV HA-E1 with truncated E1p84 by premature stop codons within the UL112/113 intron2/2', L353Stop
MCMV E1 Stop	MCMV Smith GFP with 3 point mutations introduced to create 3 premature stops in M112/113, R6Stop, G8Stop, S9Stop
MCMV E1-tev-HA	MCMV Smith GFP with C-terminal HA tagged E1p87 containing a TEV protease cleavage site
MCMV HA-tev-E1	MCMV Smith GFP with N-terminal HA tagged E1 proteins containing TEV protease cleavage sites
MCMV replacer p33 fw	MCMV Δp33 with additional insertion of an E1p33 ORF including polyA signal in the m02-m06 gene region in leading strand orientation using phosphoglycerate kinase (PGK) promoter
MCMV replacer p33 rev	MCMV Δp33 with additional insertion of an E1p33 ORF in the m02-m06 gene region in lagging strand orientation using PGK promotor and m02-m06 polyA signal
MCMV Rev E1 Stop	MCMV E1 Stop reverted back to wildtype sequence
MCMV Rev p38	MCMV Δp38 reverted back to wildtype sequence
MCMV Rev p38p36	MCMV Δp36Δp38 reverted back to wildtype sequence
MCMV Rev p87	MCMV Δp87 reverted back to wildtype sequence

MCMV SynI1	MCMV Smith GFP having the M112/113 intron 1 replaced by a synthetic intron (SynI)
MCMV Δp33	MCMV Smith GFP with deletion of the M112/113 intron 1, fusing M112/113 exon 1 and 2
MCMV Δp36	MCMV Smith GFP with inactivated E1p36 by mutation of the newly discovered splice donor D3
MCMV Δp36Δp38	MCMV Smith GFP with inactivated E1p36 and E1p38 by mutation of the shared splice acceptor A2
MCMV Δp36Δp38/p60	MCMV Smith GFP with inactivated E1p36 and E1p38 by mutation of the shared splice acceptor A2 and G164420T mutation leading to a truncated Ep87 protein, G411Stop
MCMV Δp38	MCMV Smith GFP with inactivated E1p38 by mutation of the splice donor D2
MCMV Δp87	MCMV Smith GFP with truncated E1 p87 by premature stop codons within the M112/113 intron2/2', E335Stop, G338Stop

7.3 Bacteria

name	description	growth temperature	reference / source
E.coli DH10B	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 endA1 recA1 deoR Δ(ara,leu)7697 araD139 galU GalK nupG rpsL λ-*	37 °C	Life technologies
E.coli GS1783	DH10B I *cI*857 Δ(*cro-bioA*)<>*araC*-PBAD*I-sceI*	30 °C	[128]

7.4 Plasmids

name	description	reference / source
pcDNA3	expression vector, amp^R, neo^R	Life Technologies
pCGN-pp71	pCGN plasmid carrying the HCMV gene UL82, coding for the tegument protein pp71	[167]
pCMV-SPORT6-mhnRNPA1	pCMV-SPORT6 plasmid containing the murine *hnrnpa1* gene, clone from the mammalian gene collection (MGC)	Thermo Fisher Scientific (CloneId 6308746)

pCMV-SPORT6-mhnRNPA3	pCMV-SPORT6 plasmid containing the murine *hnrnpa3* gene, clone from the MGC	Thermo Fisher Scientific (CloneId 5362605)
pCMV-SPORT6-mRBMX	pCMV-SPORT6 plasmid containing the murine *rbmx* gene, clone from the MGC	Thermo Fisher Scientific (CloneId 3599678)
pEPkan-S	template plasmid for *en passant* mutagenesis, contains I-Sce-aphA1 cassette, kan^R	N. Osterrieder, FU Berlin, Germany [128]
pMA-RQ-SynI	pMA-RQ plasmid containing an synthetic intron based on the intron within the pCI vector for subcloning and BAC mutagenesis, amp^R	Life Technologies, constructed using GeneArt service
pMSCVpuro	retroviral expression vector to generate retrovirus for transduction of eukaryotic cells, amp^R, $puro^R$	Clontech Laboratories
pReplacer	based on pBluescriptII KS+, plasmid for homologous recombination to insert genes in the m02-m06 region of MCMV, amp^R , PGK promoter	[168]
pRetroEBNA	retroviral expression vector to generate retrovirus for transduction of eukaryotic cells, amp^R, no other selection marker	T. Shenk, Princeton University, USA
pRetroGFP	based on pRetroEBNA, retroviral expression vector to generate retrovirus for GFP expression in eukaryotic cells	T. Shenk, Princeton University, USA

The following plasmids were created during the course of this work

name	description	cloning strategy
pcDNA HCMV HA-tev-E1	HCMV UL112/113 full length with N-terminal HA-tev tag cloned into pcDNA3	PCR amplification of the UL112/113 gene region from the HCMV HA-tev-E1 BAC using primer hE1 HindIII fwd and hE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.
pcDNA HCMV Δp34	HCMV UL112/113 Δp34 with N-terminal HA-tev tag cloned into pcDNA3	PCR amplification of the UL112/113 gene region from the HCMV Δp34 BAC using primer hE1 HindIII fwd and hE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.

pcDNA HCMV Δp43	HCMV UL112/113 Δp43 with N-terminal HA-tev tag cloned into pcDNA3	PCR amplification of the UL112/113 gene region from the HCMV Δp43 BAC using primer hE1 HindIII fwd and hE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.
pcDNA HCMV Δp50	HCMV UL112/113 Δp50 with N-terminal HA-tev tag cloned into pcDNA3	PCR amplification of the UL112/113 gene region from the HCMV Δp50 BAC using primer hE1 HindIII fwd and hE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.
pcDNA HCMV Δp84	HCMV UL112/113 Δp84 with N-terminal HA-tev tag cloned into pcDNA3	PCR amplification of the UL112/113 gene region from the HCMV Δp84 BAC using primer hE1 HindIII fwd and hE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.
pcDNA MCMV E1	MCMV M112/113 full length cloned into pcDNA3	PCR amplification of the M112/113 gene region from the MCMV Smith GFP BAC using primer mE1 HindIII fwd and mE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.
pcDNA MCMV Δp33	MCMV M112/113 Δp33 cloned into pcDNA3	PCR amplification of the M112/113 gene region from the MCMV Δp33 BAC using primer mE1 HindIII fwd and mE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.
pcDNA MCMV Δp36	MCMV M112/113 Δp36 cloned into pcDNA3	PCR amplification of the M112/113 gene region from the MCMV Δp36 BAC using primer mE1 HindIII fwd and mE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.
pcDNA MCMV Δp36 Δp38	MCMV M112/113 Δp36 Δp38 cloned into pcDNA3	PCR amplification of the M112/113 gene region from the MCMV Δp36 Δp38 BAC using primer mE1 HindIII fwd and mE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.

pcDNA MCMV Δp38	MCMV M112/113 Δp38 cloned into pcDNA3	PCR amplification of the M112/113 gene region from the MCMV Δp38 BAC using primer mE1 HindIII fwd and mE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.
pcDNA MCMV Δp87	MCMV M112/113 Δp87 cloned into pcDNA3	PCR amplification of the M112/113 gene region from the MCMV Δp87 BAC using primer mE1 HindIII fwd and mE1 EcoRV rev. Insertion into pcDNA3 using HindIII and EcoRV restriction sides.
pMA-RQ-SynI-EPkan-S shuttle	Shuttle vector for BAC mutagenesis based on pMA-RQ-SynI containing the I-Sce-aphA1 cassette	PCR amplification of the I-Sce-aphA1 cassette from the pEPkan-S plasmid using primer EP EcoRI fw and EP EcoRI rev. Insertion into pMA-RQ-SynI using the EcoRI restriction side.
pMSCV flag-RBMX	murine *rbmx* gene cloned into pMSCVpuro carrying an N-teminal flag tag	PCR amplification of the *rbmx* gene from pCMV-SPORT6-mRBMX using primer flag rbmx pmscv BglII fwd and rbmx mscv EcoRI rev. Insertion into pMSCVpuro using BglII and EcoRI restriction sites.
pMSCV hnRNPA1-flag	murine *hnrnpa1* gene cloned into pMSCVpuro carrying a C-teminal flag tag	PCR amplification of the *hnrnpa1* gene from pCMV-SPORT6-mhnRNPA1 using primer hnrnpa1 pmscv BglII fwd and hnrnpa1 flag pmscv EcoRI rev. Insertion into pMSCVpuro using BglII and EcoRI restriction sites.
pMSCV hnRNPA3-flag	murine *hnrnpa3* gene cloned into pMSCVpuro carrying a C-teminal flag tag	PCR amplification of the *hnrnpa3* gene from pCMV-SPORT6-mhnRNPA3 using primer hnrnpa3 pmscv BglII fwd and hnrpa3 flag pmscv EcoRI rev. Insertion into pMSCVpuro using BglII and EcoRI restriction sites.
pReplacer E1p33 lagging	MCMV E1p33 ORF cloned into pReplacer for insertion into MCMV m02-m06 region on lagging strand	PCR amplification of the E1p33 ORF from the MCMV Smith GFP BAC using primer E1 p33 replacer BamHI fwd and E1 p33 replacer NheI rev. Insertion into pReplacer using BglII and NheI restriction sides.

pReplacer E1p33 leading	MCMV E1p33 ORF cloned into pReplacer for insertion into MCMV m02-m06 region on leading strand	PCR amplification of the E1p33 ORF from the MCMV Smith GFP BAC using primer E1 p33 replacer BamHI fwd and E1 p33 replacer polyA NheI rev. Insertion into pReplacer using BamHI and NheI restriction sides.
pRetro E1p87	MCMV E1p87 ORF (lacking intron 1) inserted into pRetroEBNA	PCR amplification of the E1p87 ORF from the MCMV Δp33 BAC using primer E1p87 retro PmlI fwd and E1p87 retro NotI rev. Insertion into pRetroEBNA using PmlI and NotI restriction sites.

7.5 Primer

All Primers were purchased from Life Technologies.

name	sequence (5' to 3')	purpose
E1 p33 replacer BamHI fwd	AAGGATCCATGGCCGCGCCAGATCGACG	construction of pReplacer E1p33 leading and pReplacer E1p33 lagging
E1 p33 replacer NheI rev	AAAAAGCTAGCGCGACCAGTAAAACCCCGAT	construction of pReplacer E1p33 lagging
E1 p33 replacer polyA NheI rev	AAAAAGCTAGCTTGCTGTCCTGCCCCACCCCACCCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTCCTCATTTTATTAGGATCCTGCAGTCGACGCT	construction of pReplacer E1p33 leading
EP EcoRI fwd	AAGAATTCTAGGGATAACAGGGTAATCGATTT	construction of pMA-RQ-SynI-EPkan-S shuttle
EP EcoRI rev	AAGAATTCGCCAGTGTTACAACCAATTAACC	
hE1 EcoRV rev	AAAGATATCTTAATCGTCGAAAAACGCCG	construction of all pcDNA HCMV E1 plasmids
hE1 HindIII fwd	AAAAAGCTTATGTACCCATACGACGTCCC	
hEP A2mut fwd	AAGGAGAAAAGATTGTGCGATCTCCCCCTGGTTTCCAGCCGACTCTTGCCAGAGACGTCGTAGGGATAACAGGGTAATCGATTT	construction of HCMV Δp43
hEP A2mut rev	TTGACGACGACAGTCCCGCCCGACGTCTCTGGCAAGAGTCGGCTGGAAACCAGGGGGAGAGCCAGTGTTACAACCAATTAACC	

hEP A3mut fwd	CACTCCAGTCTCTAACTGCCGTGTTCCTCCGAATTCGCAAGAATCCGCGGCGCCTCAGCCTAGGGATAACAGGGTAATCGATTT	construction of HCMV Δp50
hEP A3mut rev	CAAAACGCGGACTGCGAGGAGGCTGAGGCGCCGCGGATTCTTGCGAATTCGGAGGAACACGCCAGTGTTACAACCAATTA	
hEP E1 stop fwd	CGTCATGGATCTCCCTACTACCGTCGTGCGAAAATAATGAACTTTTGCGAATCCTAACCGTAGGGATAACAGGGTAATCGATTT	construction of HCMV E1 stop
hEP E1 stop rev	TGACGCTCTGATGCAGGATGCGGTTAGGATTCGCAAAAGTTCATTATTTTCGCACGACGGGCCAGTGTTACAACCAATTAACC	
hEP Intron1del fwd	CAGAAGGAGCGGCGGCCGCCTCCCTCCTCCGAGAACGACGGCTCTCCTCCCCTCCGTAGGGATAACAGGGTAATCGATTT	construction of HCMV Δp34
hEP Intron1del rev	TCTTTTGTCTCTTCGCCTCCCGGAGGGGAGGAGAGCCGTCGTTCTCGGAGGAGGGAGCCAGTGTTACAACCAATTAACC	
hEP p84stop fwd	GCCGGCACCGACGGTGCGTTACTTCTACCCATTTAGCGCTGAGCGGTGGTTTCGTCGCCGTAGGGATAACAGGGTAATCGATTT	construction of HCMV Δp84
hEP p84stop rev	GACGACGGCGACGTCGACGACGGCGACGAAACCACCGCTCAGCGCTAAATGGGTAGAAGTGCCAGTGTTACAACCAATTAACC	
hEP SynI fwd	GATGAACCGGCAGAAGGAGCGGCGGCCGCCTCCCTCCTCCGAGAACGACGGTGAGTATCAAGGTTACAA	construction of HCMV SynI1
hEP SynI rev	GTGCTGTGCGGTGGTCTTTTGTCTCTTCGCCTCCCGGAGGGGAGGAGAGCCTGTGGAGAGAAAGGCAAAG	

hnrnpa1 flag pmscv EcoRI rev	TTGAATTCTTACTTGTCGTCGTCGTCCTTGTAGTCGAACCTCCTGCCACTGCCAT	construction of pMSCV hnRNPA1-flag
hnrnpa1 pmscv BglII fwd	TTTAAGCTTCTCGAGCCACCATGTCTAAGTCCGAGTCTCC	
hnrnpa3 flag pmscv EcoRI rev	TTGAATTCTTACTTGTCGTCGTCGTCCTTGTAGTCAAACCTTCTGCTACCATATC	construction of pMSCV hnRNPA3-flag
hnrnpa3 pmscv BglII fwd	TTTAAGCTTAGATCTCCACCATGGAGGTAAAACCGCCGCC	
intron 1 MCMV fwd	CAACGGTACCCTTCATAGGG	analytic PCR of MCMV intron 1 splicing
intron 1 MCMV rev	GTTCCAGATCCCTCTCCGGT	
mE1 EcoRV rev	AAGATATCTCAATTAAGATCATCGAACACATTGTCCAAGTCGA	construction of all pcDNA MCMV E1 plasmids
mE1 HindIII fwd	AAAAAGCTTATGGCCGCGCCAGATCG	
mEP A2mut fwd	AACAGAGGCTTTCTAACTGCAATCTTCCTCCGAATTCACAAGAGGACAACAACCCGCGTTTTAGGGATAACAGGGTAATCGATTT	construction of MCMV Δp36 Δp38
mEP A2mut rev	GGGAATTAAAGACGCCTTCAAAACGCGGGTTGTTGTCCTCTTGTGAATTCGGAGGAAGATTGCCAGTGTTACAACCAATTAACC′	
mEP A2rev fwd	AACAGAGGCTTTCTAACTGCAATCTTCCTCCGAATTCACAGGAGGACAACAACCCGCGTAGGGATAACAGGGTAATCGATTT	construction of MCMV Revp36p38
mEP A2rev rev	AATTAAAGACGCCTTCAAAACGCGGGTTGTTGTCCTCCTGTGAATTCGGAGGAAGATTGCCAGTGTTACAACCAATTAACC	
mEP D2mut fwd	GAACAGTTATGCTGCTAGTTCGCTGTCCGCCGTCAGCGACGGAGCGTTACCACCCTCCTCTCCTAGGGATAACAGGGTAATCGATTT	construction of MCMV Δp38
mEP D2mut rev	CTGGTGTCTCTATAGTCACCGGAGAGGAGGGTGGTAACGCTCCGTCGCTGACGGCGGACAGCGAGCCAGTGTTACAACCAATTAACC	

mEP D2rev fwd	TGAACAGTTATGCTGCTAGTTCGCTGTCCGCC GTCAGCGATGGTGCGTTACCACCCTCCTCTAG GGATAACAGGGTAATCGATTT	construction of MCMV Revp38
mEP D2rev rev	GTGTCTCTATAGTCACCGGAGAGGAGGGTGGT AACGCACCATCGCTGACGGCGGACAGCGAGCC AGTGTTACAACCAATTAACC	
mEP D3mut fwd	AGAGGGATCTGGAACCGCCTAGGAGCAGTACC ACTGTCGACGGAAATGCAGACGAAGCTGGAGC TAGGGATAACAGGGTAATCGATTT	construction of MCMV Δp36
mEP D3mut rev	AACTGTTCAAAAAGGAGACCGCTCCAGCTTCG TCTGCATTTCCGTCGACAGTGGTACTGCTCCT GCCAGTGTTACAACCAATTAACC	
mEP E1 stop fwd	CGGTGAAAAGACCTTCGTTCGGACCATGGCCG CGCCAGATTGACGCTGATAGCCCATCGTTTCG AGACGACTAGGGATAACAGGGTAATCGATTT	construction of MCMV E1 stop
mEP E1 stop rev	GTTCCGGAAGGTGAAGTAACGTCGTCTCGAAA CGATGGGCTATCAGCGTCAATCTGGCGCGGCC ATGGTCCGCCAGTGTTACAACCAATTAACC	
mEP E1 tev HA fw	ATATGGTCGACTTGGACAATGTGTTCGATGAT CTTAATGAGAACCTCTACTTCCAAGGTGGTGG TTACCCATACGATAGGGATAACAGGGTAATCG ATTT	construction of MCMV E1-tev-HA
mEP E1 tev HA rev	TACAGAGCATCATTTCTTTATCCATCTTTCAT GAGATCAAGCGTAGTCTGGGACGTCGTATGGG TAACCACCACCTTGCCAGTGTTACAACCAATT AACC	

mEP HA tev E1 fw	TTCTCGTCGCGACCGGTGAAAAGACCTTCGTTCGGACCATGTACCCATACGACGTCCCAGACTACGCTGGTGGTGGTAGGGATAACAGGGTAATCGATTT	construction of MCMV HA-tev-E1
mEP HA tev E1 rev	CGAAACGATGGGCGATCCGCGTCGATCTGGCGCGGCCATTTGGAAGTAGAGGTTCTCACCACCACCAGCGTAGTCTGGCCAGTGTTACAACCAATTAACC	
mEP Intron1del fwd	ACAAAGCAGAAGGAGCGCAGCAGGAGGCTCGAAGAGGAATGTTCTCCACGCAGCGGGGGGTAGGGATAACAGGGTAATCGATTT	construction of MCMV Δp33
mEP Intron1del rev	ATTCCTCTTCGAGCCTCCTGGTCTTCTGGCGCTTGGGCTCCCCCCCGCTGCGTGGAGAACGCCAGTGTTACAACCAATTAACC	
mEP p60 fwd	CTTTACGAGCCGCATGACGGCGGTGAAGTCGGGATCTGAGtGAGCGGGCAACACCTGTTAGGGATAACAGGGTAATCGATTT	construction of MCMV Δp36Δp38/p60
mEP p60 rev	GAAAGCCTCTGTTCCCCAGGACAGGTGTTGCCCGCTCaCTCAGATCCCGACTTCACCGGCCAGTGTTACAACCAATTAACC	
mEP p87rev fwd	CAGCGATGGTGCGTTACCACCCTCCTCTCCGGTGACTATAGAGACACCAGGGTTTAATGACTTGATAGATAGGGATAACAGGGTAATCGATTT	construction of MCMV Revp87
mEP p87rev rev	ATCGTCTTCCTCCTCGGATTCTATCAAGTCATTAAACCCTGGTGTCTCTATAGTCACCGGAGAGGAGGGCCAGTGTTACAACCAATTAACC	
mEP p87stop fwd	CAGCGATGGTGCGTTACCACCCTCCTCTCCGGTGACTATATAGACACCATGATTTAATGACTTGATAGAATCTAGGGATAACAGGGTAATCGATTT	construction of MCMV Δp87
mEP p87stop rev	TCCATCGTCTTCCTCCTCGGATTCTATCAAGTCATTAAATCATGGTGTCTATATAGTCACCGGAGAGGAGGGCCAGTGTTACAACCAATTAACC	

rbmx flag pmscv BglII fwd	TTAAGCTTAGATCTCCACCATGGACTACAAGGACGACGACGACAAGATGGTTGAAGCAGATCGCCC	construction of pMSCV flag-rbmx
rbmx pmscv EcoRI rev	TTGAATTCCTAGTATCTGCTTCTGCCTC	
seq Smith fw1	GGTTATAACAACGTCATGCA	amplification and sequencing of ML112/113 from MCMV Smith
seq Smith fw2	TGACGACCGACGTAAGAGAT	
seq Smith rev1	ACGAGGCGATCGATAACAACAGTC	
seq Smith rev2	TGCGTGGTGCCTGTA	
seq TB40/E fw1	ACGTTGCGTCGTGACGTTGT	amplification and sequencing of UL112/113 from HCMV TB40/E
seq TB40/E fw2	AAGAACGAGACGGAGCAGCA	
seq TB40/E rev1	GAACTGTGATGTCTATCTCTTTATTG	
seq TB40/E rev2	TTGTTGGAGGAAGAAGAGAC	

7.6 Antibodies

7.6.1 Primary antibodies

antigen	clone	species	used for (dilution)	reference / source
flag	M2	mouse	IB (1:1000) IP (1:300)	Sigma-Aldrich
HA	16B12	mouse	IB (1:1000)	Covance
HA	H6908	rabbit	IP (1:300)	Sigma-Aldrich
MCMV E1 (all isoforms)	serum	rabbit	IB (1:100)	J. Kerry, Eastern Virginia Medical School, USA [68]
MCMV E1 (p36, p38 and p87)	Croma 103	mouse	IB (1:1000)	S. Jonjic, University of Riejka, Croatia
MCMV gB	M55.01	mouse	IB (1:1000)	S. Jonjic, University of Riejka, Croatia
MCMV M44	3B9.22 A	mouse	IB (1:1000)	L. Loh, University of Saskatchewan, Canada
MCMV IE1	Croma 101	mouse	IB (1:1000)	S. Jonjic, University of Riejka, Croatia
β-actin	AC-74	mouse	IB (1:3000)	Sigma-Aldrich

7.6.2 Secondary antibodies

antigen	conjugate	species	used for (dilution)	reference /source
mouse Ig	HRP	goat	IB (1:3000)	DakoCytomation
mouse Ig	HRP	goat	IB (1:3000)	Jackson ImmunoResearch
rabbit Ig	HRP	goat	IB (1:3000)	Jackson ImmunoResearch

7.7 Chemicals and reagents

Standard chemicals were purchased from Sigma-Aldrich, Merck or Roth. Important chemicals and reagents are listed below.

7.7.1 Antibiotics

name	used for	concentration	reference / source
ampicillin	selection of bacteria	100 µg/ml	Roth
chloramphenicol	selection of bacteria	15 µg/ml	Roth
diphtheria toxin	selection of phoenix cells	2 µg/ml	PAA
hygromycin B	selection of phoenix cells	200 µg/ml	PAA
kanamycin	selection of bacteria	50 µg/ml	Roth
penicillin	cell culture supplement	100 U/ml	Sigma-Aldrich
puromycin	selection of transduced cells	5 µg/ml	Sigma-Aldrich
streptomycin	cell culture supplement	100 µg/ml	Sigma-Aldrich

7.7.2 Enzymes

AcTEV™ protease and buffer	Thermo Fisher Scientific
Dream Taq Green DNA polymerase and buffer	Thermo Fisher Scientific
Fast alkaline phosphatase	Thermo Fisher Scientific
Fast Digest restriction enzymes and buffer	Thermo Fisher Scientific
PRECISOR DNA polymerase and buffer	BioCat
RevertAid H Minus Reverse Transcriptase and buffer	Thermo Fisher Scientific
RNAse A	Roth
T4-DNA-ligase and buffer	Thermo Fisher Scientific

7.7.3 Molecular mass standards

O'GeneRuler™ DNA Ladder Mix	Thermo Fisher Scientific
PageRuler™ Prestained Protein Ladder	Thermo Fisher Scientific

7.7.4 SILAC reagents

L-arginine	Sigma-Aldrich
L-arginine-$^{13}C_6$	Sigma-Aldrich
L-lysine	Sigma-Aldrich
L-lysine-$^{13}C_6$	Sigma-Aldrich
L-proline	Sigma-Aldrich
stable L-glutamine	Sigma-Aldrich

7.7.5 Other reagents and chemicals

Anti-FLAG M2-agarose from mouse	Sigma-Aldrich
Anti-HA Affinity Matrix (anti-HA rat, clone 3F10)	Roche
ECL Prime Western Blotting Detection Reagent	GE Healthcare Life Science
LE Agarose	Biozym
Lumigen ECL Ultra (TMA-6)	Beckman Coulter
nitrocellulose membrane	GE Healthcare Life Science
Polybrene®	Millipore
polyethylenimine (PEI), branched	Sigma-Aldrich
PolyFect® Transfection reagent	Qiagen
protease inhibitor cocktail cOmplete™ mini, EDTA free	Roche
protein A-agarose	Roche
protein G-agarose	Roche
StrataClean Resin	Agilent Technologies
Whatman® gel blotting paper, Grade GB003	Sigma-Aldrich

7.8 Media and buffers

7.8.1 Cell culture

Dulbecco's Modified Eagle Medium (DMEM) for SILAC	Thermo Fisher Scientific
Dulbecco's Modified Eagle Medium (DMEM), high glucose	Sigma-Aldrich
Dulbecco's Phosphate Buffered Saline (PBS) (1x)	Sigma-Aldrich
fetal bovine serum, dialyzed (dFBS)	Thermo Fisher Scientific
fetal calf serum (FCS)	PAN Biotech
OptiMEM-I	Thermo Fisher Scientific
newborn calf serum (NCS)	PAN Biotech
penicillin/streptomycin (100x)	Sigma-Aldrich
trypsin-EDTA (1x)	Sigma-Aldrich

DMEM 10 % NCS	DMEM + 10 % (v/v) NCS and 1 % (v/v) penicillin/streptomycin
DMEM 10 % FCS	DMEM + 10 % (v/v) FCS and 1 % (v/v) penicillin/streptomycin
DMEM SILAC light (R0K0)	DMEM for SILAC + 10 % (v/v) dFCS and 1 % (v/v) penicillin/streptomycin additionally: 84 mg/l L-arginine, 143 mg/l L-lysine, 200 mg/l L-proline,584 mg/l stable L-glutamine sterile filtered
DMEM SILAC heavy (R6K6)	DMEM for SILAC + 10 % (v/v) dFCS and 1 % (v/v) penicillin/streptomycin additionally: 84 mg/l L-arginine-$^{13}C_6$, 143 mg/l L-lysine-$^{13}C_6$, 200 mg/l L-proline, 584 mg/l stable L-glutamine sterile filtered

7.8.2 Bacteria medium

Lysogeny broth (LB) medium (Lennox)	Roth
Lysogeny broth (LB) agar	LB medium with 15 g/l agar

7.8.3 Agarose gel electrophoresis

50x TAE buffer	2 M Tris 50 mM EDTA 5.7 % (v/v) acetic acid pH 8.0	used 1x for pouring agarose gels and as running buffer
10 x TBE buffer	990 mM Tris 40 mM EDTA 990 mM boric acid pH 8.0	used 0.5x for pouring agarose gels and as running buffer

7.8.4 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

2x sample loading buffer (SLB)	150 mM Tris 2 mM EDTA 20 % (v/v) glycerol 4 % (v/v) SDS 10 % β-mercaptoethanol bromophenol blue pH 6.8	
10 x Laemmli running buffer	250 mM Tris 1.92 M glycine 1 % (w/v) SDS	used 1x for running polyacrylamide gels
10 x TBS-T buffer	100 mM Tris 1,5 M NaCl 1 % (v/v) Tween 20 pH 7.5	used 1x for preparing antibody dilutions and washing nitrocellulose membranes
transfer buffer	50 mM Tris 150 mM NaCl 0.04 % (v/v) SDS 20 % (v/v) methanol	

7.8.5 Immunoprecipitation

RIPA lysis buffer	50 mM Tris
(sterile filtered for SILAC samples)	150 mM NaCl
	1 % (v/v) Triton X-100
	0.1 % (v/v) SDS
	1 % deoxylcholate
	pH 7.2
minimal washing buffer	50 mM Tris
(sterile filtered for SILAC samples)	150 mM NaCl
SILAC sample buffer	50 mM Tris
(sterile filtered)	150 mM NaCl
	1 % (v/v) SDS

7.8.6 DNA preparation from bacteria (“Mini” scale)

S1	50 mM Tris
	10 mM EDTA
	100 µg/ml RNAse A
	pH 8.0
S2	200 mM NaOH
	1 % (v/v) SDS
S3	2.8 M calcium acetate
	pH 5.1
TE	10 mM Tris
	1 mM EDTA
	pH 7.8

7.8.7 Kits

BCA Protein Assay Kit	Thermo Fisher Scientific
innuPREP DNA mini kit	Analytik Jena
innuPrep RNA mini kit	Analytik Jena
mi-Plasmid Miniprep Kit	Metabion
NucleoBond Gel and PCR Clean-up	Macherey-Nagel
NucleoBond Xtra Midi	Macherey-Nagel
TOPO® TA Cloning kit	Thermo Fisher Scientific
Turbo-DNA free kit	Life Technologies

8 Methods

8.1 Molecular biology methods

8.1.1 Production of electrocompetent bacteria

200 ml of warm LB medium was inoculated with 2 to 5 ml of a 10 ml overnight bacteria culture. After inoculation the 200 ml cultures were grown in a shaking incubator HT (Infors) until reaching the exponential growth phase at an OD_{600} of 0.5 to 0.6. Cultures of strain DH10B were cooled down on ice immediately, while cultures of the GS1783 or SW102 strains were incubated at 42 °C in order to induce recombinase expression before cooling down on ice. Bacteria were pelleted by centrifugation at 4 °C using 6000 x *g* for 10 minutes in a RC-6 centrifuge (Sorvall). The resulting pellets were washed twice with 100 ml cold autoclaved deionized water and once with 50 ml cold autoclaved 10 % (v/v) Glycerol before being resuspended in 2 ml autoclaved 10 % (v/v) Glycerol. Aliquots of 100 µl were made and shock frozen at -80 °C.

8.1.2 Transformation and storage of bacteria

Electrocompetent bacteria were used for transformation by electroporation. 50 µl of the aliquoted bacteria were mixed with either 200 ng of PCR amplified linear DNA, 1 ng of supercoiled plasmid or 2.5 µl of ligation mixes. After 10 minutes of incubation on ice the bacteria-DNA-suspensions were transferred into 2 mm electroporation cuvettes and pulsed using the settings of 2500 V, 25 µF and 200 Ω with the Gene Pulser XCell (BioRad). After pulsing, 900 µl of warm LB medium was added to the bacteria. This was then incubated on a Thermomixer comfort 5355 (Eppendorf) for 1 hour before being plated on LB agar for overnight incubation in a bacteria incubator (IPP400, Memmert).

For long term storage 700 µl of an overnight bacteria culture was mixed with 300 µl of autoclaved 80 % (v/v) glycerol and frozen at -80 °C.

8.1.3 Isolation of DNA from bacteria

8.1.3.1 Small scale plasmid DNA preparation (“Plasmid Mini Prep”)

Plasmid DNA was isolated from 5 ml overnight cultures using the Mi-Plasmid MiniPrep Kit according to the manufacturer’s guidelines. Plasmid DNA was resolved in 50 µl elution buffer.

8.1.3.2 Small scale BAC DNA preparation ("BAC Mini Prep")

BAC DNA was isolated from 5 ml overnight cultures using alkaline lysis and isopropanol precipitation [169]. 2 ml of the bacteria culture was pelleted at 6000 x *g* for 5 minutes using a benchtop centrifuge 5415R (Eppendorf). The supernatants were then discarded and pellets were resolved in 300 µl of buffer S1. Cell lysis was initiated by adding 300 µl of buffer S2, and cells were allowed to lyse for 5 minutes before lysis was stopped by adding 300 µl of buffer S3. Neutralization of the lysis reaction leads to precipitation of both bacterial chromosomal DNA as well as proteins, of which the proteins were first pelleted by centrifugation using 15000 x *g* for 10 minutes at 4 °C. The supernatants were subsequently transferred to new tubes and BAC DNA was precipitated using 600 µl isopropanol and centrifugation using 15000 x *g* for 30 minutes at 4 °C. Supernatants were then discarded and the DNA pellets were washed using 500 µl of 70 % (v/v) ethanol. The ethanol was removed and pellets were dried at room temperature until no residual ethanol was visible. The BAC DNA was finally rehydrated using 50 µl of TE buffer.

8.1.3.3 Medium scale BAC and plasmid DNA preparation ("BAC/Plasmid Midi Prep")

BAC or Plasmid Midi Preps were done using 200 ml bacteria cultures. The DNA was isolated using the NucleoBond Midi Xtra Kit according to the manufacturer's guidelines (for BAC DNA the low copy protocol was used). BAC was eluted using 150 µl TE buffer, while Plasmid DNA was eluted using 500 µl TE buffer.

8.1.4 Polymerase Chain Reaction (PCR)

PCRs were performed using either Precisor Polymerase or DreamTaq Polymerase according to the manufacturer's guidelines. Dream Taq was used for analytical PCRs and TOPO TA cloning while Precisor was used for all other cloning purposes.

8.1.5 Restriction digestion of DNA

DNA was digested using FastDigest (FD) restriction enzymes following the manufacturer's instructions. Plasmids which were digested for use as a vector in ligation reactions were additionally treated with Fast Akaline Phosphatase (AP) according to the manufacturer's recommendations. For analytical plasmid digestion 1 µg DNA was used whereas for preparative plasmid digestion 2 µg DNA was used. Analytical BAC digestion was performed for 1h at 37 °C using between 0.5 and 2 µg of BAC Midi Prep DNA or 17 µl of BAC Mini Prep DNA.

8.1.6 Agarose gel electrophoresis

PCR products and plasmid fragments were run on 1 % (w/v) agarose TAE gels for 1 hour at 120 V. Digested BACs were run on 0.6 % (w/v) agarose TBE gels overnight at 60 V. In both cases gels contained 0.5 µg/ml ethidium bromide and O'GeneRuler was used as a size ladder. DNA bands were visualized by UV light using a GelDoc XR+ (BioRad).

8.1.7 Purification of DNA fragments

DNA bands of interest were cut out from agarose gels on an UV transilluminator (Vilber) using a clean scalpel. DNA was then purified from the gel pieces using a NucleoSpin Gel and PCR clean up kit according to the manufacturer's guidelines. The same kit was also used for direct purification of PCR products or digested plasmids.
Concentration (OD_{260}) and purity (OD_{260}/OD_{280}) of DNA was measured using a NanoDrop-1000 (Peqlab) photometer.
DNA was stored at 4 °C for short-term periods (no longer than 4 weeks) or at – 80°C for long-term periods.

8.1.8 DNA ligation

Ligations were performed using T4-DNA-Ligase. Vector and insert were mixed in a molecular ratio of 1:5 to 1:7 in a total volume of 20 µl. Ligation reactions were carried out overnight at 16 °C.

8.1.9 Isolation of total RNA

Total RNA was isolated from 1 x 10^6 (CMV infected) eukaryotic cells using the innuPREP RNA Mini Kit according to the manufacturer's guidelines. Total RNA was eluted using 44 µl RNase-free water and the remaining DNA was digested using TURBO-DNA-free kit following the manufacturer's recommendation. Concentration (OD_{260}) of RNA was measured using a NanoDrop-1000 (Peqlab) photometer.
Total RNA was stored at – 80 °C.

8.1.10 Isolation of viral DNA

Viral DNA was isolated from the supernatant of 5 x 10^6 (CMV infected) cells. Virus particles were pelleted by centrifugation at 4 °C using 15000 x *g* for 3 hours before using innuPREP DNA Mini Kit to extract DNA from the pellet. The kit was used according to the manufacturer's guidelines, with the resulting DNA eluted using 60 µl of nuclease-free

water pre-heated to 70 °C.
Concentration (OD_{260}) and purity (OD_{260}/OD_{280}) of DNA was measured using a NanoDrop-1000 (Peqlab) photometer.
DNA was stored at 4 °C for short time periods (no longer than 4 weeks) or at – 80°C for long term periods.

8.1.11 DNA Sequencing

All DNA sequencing reactions were performed by GATC Biotech, Konstanz, Germany using their LightRun NXP service. Plasmid DNA and PCR products were sequenced directly by GATC. For sequencing of BAC or viral DNA the region of interest was amplified by PCR before being sent for sequencing.

8.1.12 Synthesis of cDNA

For synthesis of cDNA, 2 µg of total RNA was reverse transcribed using RevertAid H Minus Reverse Transcriptase and oligo[dT]18 according to the manufacturer's guidelines. The reverse transcriptase was inactivated by a final heating step using 70 °C for 10 minutes. Synthesized cDNA was stored at -20 °C.

8.1.13 Whole transcriptome shotgun sequencing (RNA-Seq)

RNA-Seq was performed in the HPI core facility Next Generation Sequencing by Daniela Indenbirken.
Briefly: mRNA was extracted from total RNA using the NEBNext Poly(A) mRNA Magnetic Isolation module (New England Biolabs) and RNA-seq libraries were generated using the NEXTflex Directional RNA-Seq Kit (dUTP based) for Illumina (Bio Scientific) according to the manufacturer´s instructions. Proper fragment size distribution and quality of the libraries were verified on a BioAnalyzer High Sensitivity Chip (Agilent). Diluted libraries were multiplex-sequenced on the Illumina HiSeq 2500 instrument (2 × 125 bp run) with 61 and 70 million reads per sample.

8.1.14 TOPO TA Cloning

To directly clone linear PCR fragments into plasmids the TOPO TA Cloning kit was used according to the manufacturer's guidelines. Fragments for cloning were amplified using Dream Taq Polymerase and subsequently gel purified. Plasmids derived from cloning were directly sequenced.

8.1.15 *En passant* BAC mutagenesis

Mutation of BACs using *en passant* was performed essentially as described by Tischer and colleagues [128]. In brief: linear DNA fragments containing the I-SceI-aphAI-cassette as well as a duplicate were generated using PCR and either the pEP-Kan-S plasmid or shuttle plasmids were used as a template. After purification 200 ng of the PCR products were transformed in GS1783 carrying either the MCMV Smith GFP or HCMV HA-tev-E1 BAC. Transformed bacteria were then spread on LB agar plates containing chloramphenicol and kanamycin and incubated overnight at 30 °C. Resulting bacterial clones were checked via HindIII restriction digestion, analytical PCR and sequencing, with positive clones used for the second recombination procedure. This procedure requires expression of recombinases such as I-SceI, which was induced by incubation at 42 °C and the addition of 2 % (w/v) L-arabinose. Recombination bacteria were then plated on LB agar containing 10 % L-arabinose and chloramphenicol. Resulting bacterial clones were checked for loss of kanamycin resistance as well by HindIII restriction digestion, analytical PCR and sequencing. Positive clones were outgrown in 200 ml liquid culture for BAC Midi Prep. For the construction of MCMV replacer p33 fw and MCMV replacer p33 rev no second recombination was needed.

8.2 Cell biology and virology methods

8.2.1 Cell culture and cell counting

All cells were incubated in a Hera Cell CO^2 incubator (Heraeus) at 37 °C, 80 % relative humidity and 5 % CO_2 on 145 mm dishes. All work was done using a sterile bench (HeraSafe, Heraeus). DMEM 10 % NCS was used as culture media for NIH-3T3 cells while DMEM 10% FCS was used for all other cells. Cells were split when reaching an approximately 90 % confluent monolayer. In this procedure cells were first washed with PBS and then trypsinized. The trypsin was then neutralized with double the volume of NCS or FCS supplemented medium, after which cells were split 1:3 to 1:10 on new 145 mm dishes. If cell numbers needed to be determined 10 µl of the cell suspension were analyzed using an automated cell counter (TC10, BioRad).

8.2.2 Freezing and thawing of cells

For freezing, a 90 % confluent 145 mm dish of cells was trypsinized and neutralized using 10 ml of NCS or FCS supplemented medium. Cells were then pelleted using a centrifuge (5810R, Eppendorf) at 200 x *g* for 7 minutes. Supernatant was discarded and pellet was washed in PBS. After the cells were pelleted again they were resuspended in 4 ml freeze mix and equally distributed in 4 cryotubes, after which the cells were frozen at -80 °C. Frozen cells were kept in liquid nitrogen for long-term storage.
Cells were thawed using a 37 °C waterbath and immediately transferred to a 100 mm dish and supplemented with 10 ml of DMEM 10 % NCS or DMEM 10 % FCS.

8.2.3 Transfection of plasmid DNA

Phoenix cells were transfected using PEI. For transfection 3 x 10^6 cells were seeded on 100 mm dishes. After overnight incubation those cells were transfected using 8 µg of plasmid DNA (pMSCV or pRetro). In order to do this the plasmid DNA was first diluted in 500 µl DMEM without supplements, in parallel 32 µl PEI was also diluted in 500 µl DMEM without supplements. After 10 minutes of incubation both mixes were combined and incubated for an additional 20 minutes before being added to the cells. The medium of PEI-transfected cells was changed after 12 hours.
NIH-3T3 and HEK-293A cells were transfected using Polyfect. For transfection, 5 x 10^5 cells were seeded on 6-well plates. After overnight incubation, cells were transfected using 1.5 µg of plasmid DNA. In order to do this, plasmid DNA was first diluted in 200 µl DMEM without supplements. After 2 minutes of incubation time 10 µl of Polyfect was added and DNA-Polyfect mixes were incubated for an additional 20 minutes. After the final incubation the DNA-Polyfect mixes were diluted in 500 µl DMEM + 10% FCS/NCS and added to the cells.

8.2.4 Transfection of BAC DNA

BAC DNA was transfected in eukaryotic cells in order to reconstitute CMV. For MCMV reconstitution NIH-3T3 cells were transfected using Polyfect. 1.5 x 10^5 cells were seeded on 6-well plates and transfected using 3 µg of plasmid DNA following overnight incubation. In order to do this, plasmid DNA was diluted in 200 µl DMEM without supplements. After 2 minutes of incubation time 32 µl of Polyfect was added and DNA-Polyfect mixes were incubated for an additional 20 minutes. After the final incubation

the DNA-Polyfect mixes were diluted in 500 µl DMEM + 10% NCS and added to the cells.

For HCMV reconstitution MRC-5 cells were transfected by electroporation. For transfection 1 x10^7 cells were pelleted after trypsin neutralization by centrifugation at 180 g for 8 minutes. Cells were subsequently washed using 10 ml of OptiMEM-I and pelleted again. The cell pellet was then resuspended in 500 µl OptiMEM-I. Meanwhile 2 µg pCGN-pp71 plasmid DNA was diluted in 80 µl OptiMEM-I and mixed with 10 µg of BAC DNA. The DNA mix was then added to the cell suspension and transferred to 4 mm electroporation cuvettes. Electroporation was done using a Gene Pulser XCell with 220 V and 950 µF. After pulsing, 500 µl of OptiMEM-I was carefully added to the transfected cells. Floating debris was aspirated before the cell suspension was added to 30 % confluent MRC-5 seeded on 100 mm dishes. After overnight incubation old media was replaced with 12 ml of fresh DMEM + 10 % FCS.

Reconstitution of MCMV and HCMV was monitored and documented by detection of CPEs and GFP-positive foci using an inverted fluorescence microscope Axiovert 40 CFL (Zeiss).

8.2.5 CMV infection of cells

Cell were infected with CMV using different multiplicities of infection (MOI) based on the tissue culture infection dose 50 per milliliter ($TCID_{50}$/ml) of a virus stock. The $TCID_{50}$/ml describes the amount of a virus dilution needed to infect 50 % of all cells. To determine the volume of virus stock needed to infect cells at a given MOI the following equation was used:

$$\frac{(number\ of\ cells)\ *\ MOI}{TCID_{50}/ml} = volume\ of\ virus\ stock\ in\ ml$$

The required volume of CMV stock was diluted in 200 ml of DMEM + 10 % FCS/NCS and directly added to the cells. If not stated differently all experiments were done with centrifugal enhancement of the infection at 37 °C using 1000 x g for 30 minutes [170].

8.2.6 Preparation of CMV stocks

CMV stocks were harvested from highly infected 10.1 or NIH-3T3 cells for MCMV or highly infected MRC-5 cells for HCMV. If cells were transfected with BAC DNA they were grown until two 145 mm dishes were fully infected. The supernatant of those cells was then equally distributed on ten 145 mm dishes containing 5 x 10^6 cells per plate (for

MCMV stocks) or fifteen 145 mm dishes containing 3 x 10^6 cells per plate (for HCMV stocks). Cells were incubated until fully infected, where the total supernatant was collected and combined. Cell debris was removed by centrifugation at 4 °C using 6000 x *g* for 15 minutes. The supernatant was then transferred to fresh tubes and centrifuged at 4 °C using 15000 x *g* for 4 h. The supernatant was then discarded and the resulting virus pellet was resuspended in 2 ml DMEM + 10% FCS/NCS. This virus stock suspension was aliquoted and stored at – 80 °C.

If a CMV stock was made from a preexisting virus stock, ten 145 mm dishes containing 5 x 10^6 10.1 cells per plate (for MCMV stocks) or fifteen 145 mm dishes containing 3 x10^6 MRC-5 cells per plate (for HCMV stocks) were seeded. These cells were infected the following day using an MOI of 0.02. Infected cells were incubated until fully infected and virus stock was harvested as described above.

8.2.7 Titration of CMV-containing solutions

To determine the virus concentration in CMV stocks or in cell culture supernatant the $TCID_{50}$/ml method was used. For MCMV, 2000 NIH-3T3 cells were seeded in each well of a 96-well plate, while for HCMV, 1000 MRC-5 cells were seeded in each well of a 96-well plate. Serial dilutions of CMV stocks or supernatant were prepared the following day; dilution ranges for CMV stock ranged from 1:10^2 to 1:10^9 and for supernatant ranged from 1:10^1 to 1:10^8. Each dilution was added to one entire row of wells for two 96-well plates. Plates were then centrifuged at 37 °C using 1000 x *g* for 30 minutes for the centrifugal enhancement titer before being incubated or incubated directly for the titer without centrifugal enhancement. After 6 days (MCMV) or 14 days (HCMV) of incubation the viral titer was determined by counting the number of infected wells per row using the Spearman-Karber method [171].

8.2.8 Viral replication kinetics

For replication kinetics cells were seeded in 6-well dishes. Infections were done without centrifugal enhancement in triplicates using 1 x 10^5 NIH-3T3, 8 x 10^4 SVEC4-10 or 3 x 10^5 RAW267.4 with an MOI of 0.02, 0.1 or 0.5 respectively. Virus dilution used for initial infection was titrated to determine input titer (see section 8.2.7). At 4 hours post-infection cells were washed twice with PBS and 2 ml medium was added. Supernatants were subsequently harvested at different times post-infection and 2 ml of medium was added to the infected cells. The collected supernatants were subjected to virus titration.

8.2.9 Production of retrovirus

Retrovirus was harvested from Phoenix cells transfected either with pRetro or pMSCV derived plasmids. Transfection of Phoenix cells is described in section 8.2.3. Virus containing supernatant was collected 48 h and 72 h hours after transfection, sterile filtered using a 0.45 µm filter and stored at -80°C.

8.2.10 Retroviral transduction of cells

For transduction 3 x10^4 NIH-3T3 cells were seeded in 2 wells of a 12-well plate. Old media was replaced the following day with 1.5 ml of retrovirus containing supernatant supplemented with 5 µg/µl Polybrene. Cells were then centrifuged at 37 °C using 1000 x *g* for 30 minutes. After 6 hours the virus-containing media was replaced with DMEM + 10 % NCS. The transduction procedure was repeated following overnight incubation.

If cells were transduced with pRetro-derived retrovirus they were used up to 4 weeks post transduction. If cells were transduced with pMSCV derived retrovirus they were selected using 1 µg/µl puromycin. Selective pressure was kept while those cells were in culture. Puromycin was only removed before infection or transfection experiments.

8.2.11 Stable isotope labeling by amino acids in cell culture (SILAC)

For SILAC, NIH-3T3 cells were cultivated in SILAC medium, supplemented with 10% dialyzed FCS in the presence of light arginine and lysine or heavy arginine and lysine (+ 6.0201 Da). To ensure full incorporation of the amino acids (>97 %), the labeling efficiency was checked by bottom-up LC-MS for passage 1, 3 and 5 after switching to SILAC media.

Cells were lysed in 1% SDS using ultrasonication followed by centrifugation at 4° C at 20,000 g to obtain a clear soluble protein fraction.

The next steps were performed by Stefan Loroch at the Leibniz Institute for Analytical Sciences (ISAS) in Dortmund as follows: Proteins were reduced using 10 mM dithiothreitol for 30 minutes at 56 °C and reduced cysteines were carbamidomethylated using 20 mM iodoacetamid for 30 minutes in the dark. The protein was precipitated by adding 9 volumes of ice-cold ethanol followed by incubation at -40 °C for 1 h. After centrifugation for 30 minutes at 20000 x *g* and 4 °C, the resulting pellet was resolubilized in 6 M Guanidinium chloride (GuHCL). The protein concentration was

determined using a bicinchoninic acid assay (Thermo Scientific). For tryptic digestion, the protein sample was diluted 1:20 in 50 mM Ammonium bicarbonate (ABC), 1 mM $CaCl_2$ and trypsin was added 1:30 (trypsin:sample). Digests were incubated for 16 h at 37 °C with slight agitation. The reaction was stopped by adding trifluoroacetic acid (TFA) to a final concentration of 1%. Approximately 1 µg of each digest was subsequently analyzed by LC-MS (see below). Raw-files were searched against the Uniprot database (taxonomy-mouse) with oxidation of methionine, heavy-labeled lysine and heavy-labeled arginine as variable modification. Carbamidomethylation of cysteines was set as static modification. To determine the level of label-incorporation, the percentages of peptide-spectrum matches (PSM) derived from heavy-labeled peptides were calculated. The label incorporation was found to be 93.8 %, 98.6% and 98.4 % for passage 1, 3 and 5 respectively.

8.3 Protein biochemistry methods

8.3.1 Cell lysis for immunoblotting and immunoprecipitation

Infected or transfected cells used in immunoblotting (IB also called western blot, WB) were lysed directly in 2x Laemmli buffer. This required cells washed with PBS and subsequently lysed using 200 µl of 2x Laemmli buffer per well of a 6-well plate. Cell lysates were either used directly or stored at – 20 °C.

Samples used in immunoprecipitation (IP) experiments were lysed in RIPA buffer. In this method cells were washed twice with PBS and lysed using 200 µl of RIPA buffer supplemented with cOmplete mini EDTA free protease inhibitor tablets. Complete removal of the cells from the plates was achieved through use of a cell scraper. Lysates of one entire 6-well plate were combined for a single IP sample. In order to ensure complete lysis, samples were incubated on a rotating platform at 4 °C for 20 minutes. Following this, protein concentration of the sample was measured using the BCA Protein assay kit, which was measured in 96-well plates using a BSA standard curve. 5 µl of protein sample was mixed with 5 µl of PBS in duplicate, with the BSA standard curve - using 0.125, 0.25, 0.5, 1.0 und 2.0 mg/ml BSA in PBS - also done in duplicate. Standards and samples were supplemented with 100 µl of a 50:1 mix of the BCA solutions A and B. After 20 minutes of incubation at 37 °C, absorbance at 562 nm was measured. The BSA standard curve was used to calculate the concentration of the IP samples.

8.3.2 SDS polyacrylamide gel electrophoresis (SDS-PAGE) and western blot

SDS-PAGE was carried out according to Laemmli [172]. Protein lysates were first denatured by heating to 95 °C for 10 minutes. Samples were then concentrated on 4 % stacking gel at 60 V for 15 minutes and separated in a 10 % resolving gel at 120 V for 40 minutes using the MiniPROTEAN Tetra Cell-System (BioRad). After separation the proteins were transferred on a nitrocellulose membrane (Hybond ECL, GE Healthcare) by semi-dry sandwich blotting using 25 V for 60 minutes in a Transblot Semi-dry Transfer Cell (BioRad). After transfer, the membranes were blocked using 5 % (w/v) milk powder in TBS-T for 1 hour. Membranes were then incubated in primary antibodies diluted according to section 7.6.1 in 5 % (w/v) milk powder in TBS-T overnight at 4 °C. Horseradish peroxidase (HRP)- coupled secondary antibodies were diluted according to section 7.6.2 in 5 % (w/v) milk powder in TBS-T and added to the membranes after three 10 minute washes in TBS-T. Following 1 hour incubation in secondary antibody the membranes were again washed three times for 10 minutes.

In order to detect the attached protein the membranes were covered with ECL Western Blotting Detection Reagents containing 10 % (v/v) Lumigen TMA-6 and chemiluminescence was detected either directly on the membrane with the detector Fusion SL-4 3500WL Molecular Imaging (Peqlab) or using RP new X-ray films (CEA) and a Structurix M ECO film developer (GE Measurement & Control).

8.3.3 Immunoprecipitation for western blot

Lysates from MCMV infected NIH-3T3 cells (MOI of 5) were used for IP. After RIPA lysis of the cells and measurement of the respective protein concentrations, 1.5 mg of each whole cell lysate were used for IP and the leftover was kept as lysate controls.

For IP, samples were precleared using a mix of 25 µl protein A agarose beads (PAA) and 25 µl protein G agarose beads (PGA) using a rotating platform for 1 h at 4 °C. Beads were pelleted by centrifugation for 1 minute at 15000 x *g*. The supernatants were then transferred to new tubes and affinity purification of the tagged target protein was done using 75 µl Anti-HA Affinity Matrix or anti-FLAG M2-Agarose from mouse. After 4 hours beads were pelleted by centrifugation for 1 minute at 15000 x *g* and the supernatants were discarded. The beads were washed 3 times using minimal washing buffer. The washing buffer was completely removed, the beads were resuspended in 120 µl 2x Laemmli buffer and proteins were denatured by heating at 95 ° C for 10 minutes.

The beads were then pelleted by centrifugation for 3 minutes at 15000 x *g* and the supernatants were either subjected to immunoblotting or stored at – 20 °C.

8.3.4 Immunoprecipitation for mass spectrometry analysis

IP for mass spectrometry analysis made use of lysates from heavy and light labeled MCMV infected NIH-3T3 cells (MOI of 5). After RIPA lysis of the cells and measurement of protein concentration, 1.5 mg of each whole cell lysate was used for IP and the leftover was kept as lysate controls.

Samples were precleared using a mix of 25 µl PAA and 25 µl PGA using a rotating platform for 1 h at 4 °C. Beads were afterwards pelleted by centrifugation for 1 minute at 15000 x *g*. The supernatants were transferred to new tubes and affinity purification of HA-tagged target protein was carried out using 75 µl Anti-HA Affinity Matrix. After 4 hours the beads were pelleted by centrifugation for 1 minute at 15000 x *g* and the supernatants were discarded. The beads were then washed two times using minimal washing buffer and once using AcTEV Protease buffer. The AcTEV Protease buffer was completely removed and beads were resuspended in 1 ml AcTEV Protease buffer containing AcTEV Protease, following which protease digestion was done overnight on a rotating platform at 4 °C. The beads were then pelleted by centrifugation for 2 minute at 15000 g and the supernatants were transferred to new tubes. Proteins were subsequently pulled down from the supernatants by adding 50 µl StrataClean Resin beads and rotating incubation for 1 h at 4 °C. The supernatants were then completely removed and the beads were resuspended in 60 µl SILAC sample buffer, following which the proteins were denatured for 10 minutes by heating to 95 ° C. 10 µl of each of the final lysates was kept for immunoblotting while the other 50 µl was either directly sent for mass spectrometry analysis or stored at – 80 °C.

8.3.5 Mass spectrometric analysis of immuneprecipitates

Mass spectrometry analysis of the heavy and light amino acid labeled NIH-3T3 and IP samples were carried out in the Leibniz Institute for Analytical Sciences (ISAS) in Dortmund, Germany by Stefan Loroch.

Procedure was as follows: Heavy and Light samples were mixed 1:1. Protein disulfides were reduced using 10 mM dithiotreitol for 30 minutes at 56 °C and reduced cysteines were carbamidomethylated using 20 mM iodoacetamide for 30 minutes in the dark. Proteins were precipitated by adding nine-times the volume of ice-cold ethanol (-40 °C),

thorough mixing and incubation for 1 h at -40 °C. The precipitate was spun down and washed with 20 µl ice-cold acetone followed by another centrifugation. The precipitated protein was reconstituted in 10 µl 2 M GuHCl and proteins were digested by adding 40 µl 50 mM ABC, 5 µl 10 mM $CaCl_2$ and 5 µl trypsin with a concentration of 4 ng/µl. After 16 h of incubation at 37 °C the digest was halted by adding TFA to a final concentration of 1%. Samples were dried *in vacuo*, reconstituted in 0.1 % TFA and up to 50% of each digest was subjected to nanoLC-MS/MS analysis.

Nano-LC-MS/MS was performed using either an LTQ Orbitrap Velos or a Q-Exactive mass spectrometer (MS) online coupled to a U3000 nanoHPLC (both Thermo Scientific) equipped with a C18 main- and C18 precolumn. Samples were loaded onto the precolumn in 0.1% TFA with a flowrate of 10-20 µl /min. After 10 minutes the precolumn was switched in-line with the main column and peptides were separated using a linear acetonitrile gradient (2-35 %) in 0.1 % formic acid at a flowrate of 270 nl/min. Peptides were introduced to the MS using a nanoESI source operated in positive ion mode. Data-dependent acquisition (topN) was performed with a high-resolution survey scan in the orbitrap and isolation and fragmentation of the 10 or 15 most intense peptide ions. Fragmentation was done using either higher-energy collision-induced dissociation (Q-Exactive) or collision-induced dissociation (LTQ Orbitrap Velos). Fragment spectra were acquired at lower resolution in the orbitrap (Q-Exactive) or in the linear ion trap (LTQ Orbitrap Velos) and fragmented ion masses were excluded for at least 10 sec from re-fragmentation (dynamic exclusion).

8.4 Data analysis

8.4.1 RNA-Seq

In silico splice site prediction was performed using Hbond scoring (http://www.uni-duesseldorf.de/rna/html/hbond_score.php).

The analysis of the RNA-Seq data was done by Malik Alawi from the Bioinformatics Core at the University Medical Center Hamburg-Eppendorf, Germany.

In summary: Trimmomatic [173] was employed to remove adapters and low-quality (phred quality score < 15) bases from the 3' ends of sequencing reads. A split read alignment was performed by aligning the trimmed reads to the MCMV Smith NCBI reference sequence (NC_004065.1) or the MCMV K181 Genbank entry (AM886412.1) using the software segemehl [174]. The positions of splice junctions were obtained from

the alignments following the procedure described in segemehl's manual.
All genomic positions refer to the MCMV Smith or MCMV K181 reference sequence (GenBank accession numbers NC_004065 and AM886412 respectively). For HCMV the TB40/E GenBank sequence EF999921 was used (Splice junctions are depicted with the last position of the upstream exon and the first position of the downstream exon separated by a circumflex (^).

8.4.2 Mass spectrometry

Data from MS analysis was analyzed in the Leibniz Institute for Analytical Sciences (ISAS) in Dortmund, Germany by Stefan Loroch.
Briefly: Raw-files were searched using Mascot v2.4.0 (Matrix Science) implemented in Proteome Discoverer (PD) 1.4 (Thermo Fisher Scientific). Decoy searches were performed against the Uniprot mouse database combined with the Uniprot/Trembl entries for MCMV (www.uniprot.org). Four additional sequences for E1p33, E1p36, E1p38 and E1p87 were manually added (16603 entries in total - as of 30.07.2012). Search tolerances were set to 10 ppm for precursors and 0.02 Da (Q-Exactive) or 0.5 Da (LTQ Orbitrap Velos) for fragment ions. Carbamidomethylation of cysteines was set as static, oxidation of methionine and acetylation of protein N-terminii as variable modification. The enzyme selected was trypsin (C-terminal cleavage of lysine and arginine if no proline follows) with a maximum of two missed cleavages. Quantification was conducted using the precursor ion quantifier node of PD with a 2 ppm quantification window. Results were filtered using a 1 % false discovery rate on peptide spectra level (PD) and proteins were only considered if quantified by at least 2 unique peptides. Lists were exported to Microsoft Excel and further processed either in Excel or R v3.0.3. For quantification, Log2 ratios were calculated and median centered. Proteins were considered as potentially regulated if the ratio was higher than two times the standard deviation over all protein ratios (FIG 26).

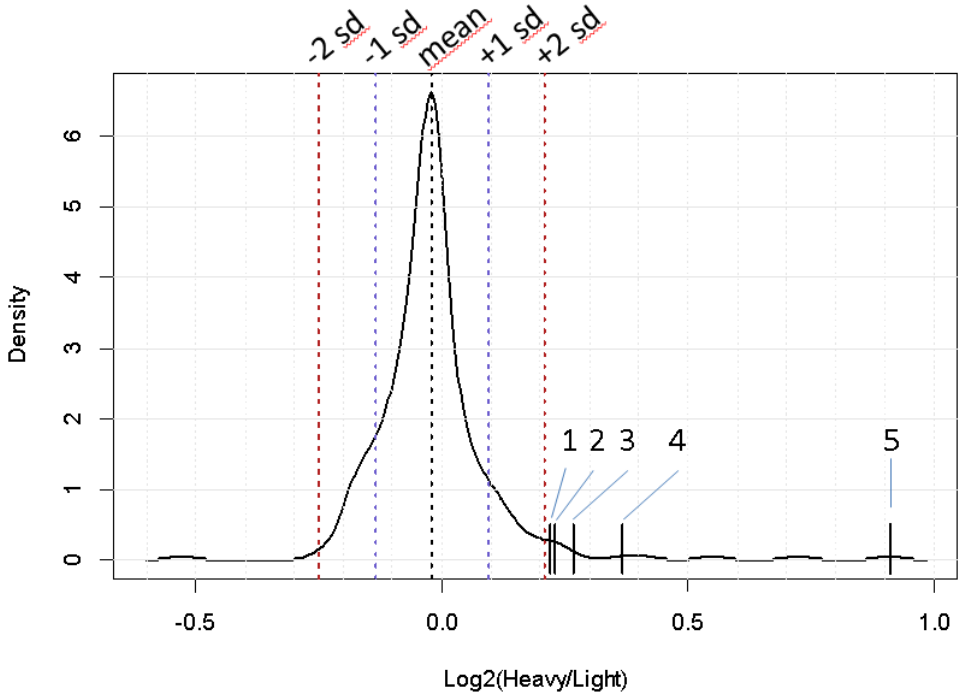


FIG 26 Example of a SILAC-based IP of E1p87-HA. 413 proteins were identified. Proteins with a heavy-to-light ratio (log2) above the mean + two standard deviations (sd) were considered as potentially regulated. The plot shows a density distribution of log2 ratios with potentially regulated proteins marked by black bars. 1: Heterogeneous nuclear ribonucleoproteins A2/B1, 2: Heterogeneous nuclear ribonucleoprotein M, 3: Heterogeneous nuclear ribonucleoprotein A1, 4: Heterogeneous nuclear ribonucleoprotein A3, 5: RNA binding motif protein, X-linked-like-1. Figure by Stefan Loroch.

9 References

1. McGeoch, D.J., F.J. Rixon, and A.J. Davison, *Topics in herpesvirus genomics and evolution.* Virus Res, 2006. **117**(1): p. 90-104.
2. Davison, A.J., *Evolution of sexually transmitted and sexually transmissible human herpesviruses.* Ann N Y Acad Sci, 2011. **1230**: p. E37-49.
3. Weller, T.H., *Review. Cytomegaloviruses: the difficult years.* J Infect Dis, 1970. **122**(6): p. 532-9.
4. Ho, M., *The history of cytomegalovirus and its diseases.* Med Microbiol Immunol, 2008. **197**(2): p. 65-73.
5. Cannon, M.J., T.B. Hyde, and D.S. Schmid, *Review of cytomegalovirus shedding in bodily fluids and relevance to congenital cytomegalovirus infection.* Rev Med Virol, 2011. **21**(4): p. 240-55.
6. Stern, H., *Intrauterine and perinatal cytomegalovirus infections.* J Antimicrob Chemother, 1979. **5 Suppl A**: p. 81-5.
7. Kurath, S., et al., *Transmission of cytomegalovirus via breast milk to the prematurely born infant: a systematic review.* Clin Microbiol Infect, 2010. **16**(8): p. 1172-8.
8. Cannon, M.J., D.S. Schmid, and T.B. Hyde, *Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection.* Rev Med Virol, 2010. **20**(4): p. 202-13.
9. GIDEON Informatics, I. and S. Berger, *Cytomegalovirus Infection: Global Status*. 2015: GIDEON Informatics, Incorporated.
10. Voigt, S., A.S. Rosario, and A. Mankertz, *Cytomegalovirus seroprevalence among children and adolescents in Germany: Data from the KiGGS interview and examination survey, 2003-2006.* Open Forum Infectious Diseases, 2015.
11. Britt, W., *Manifestations of human cytomegalovirus infection: proposed mechanisms of acute and chronic disease.* Curr Top Microbiol Immunol, 2008. **325**: p. 417-70.
12. Popovic, M., et al., *Human cytomegalovirus infection and atherothrombosis.* J Thromb Thrombolysis, 2012. **33**(2): p. 160-72.
13. Lancini, D., et al., *Cytomegalovirus disease in immunocompetent adults.* Med J Aust, 2014. **201**(10): p. 578-80.
14. Danner, S.A., *Management of cytomegalovirus disease.* AIDS, 1995. **9 Suppl 2**: p. S3-S8.
15. Deayton, J.R., et al., *Importance of cytomegalovirus viraemia in risk of disease progression and death in HIV-infected patients receiving highly active antiretroviral therapy.* Lancet, 2004. **363**(9427): p. 2116-21.
16. Lumbreras, C., et al., *Cytomegalovirus infection in solid organ transplant recipients.* Clin Microbiol Infect, 2014. **20 Suppl 7**: p. 19-26.
17. Toupance, O., et al., *Cytomegalovirus-related disease and risk of acute rejection in renal transplant recipients: a cohort study with case-control analyses.* Transpl Int, 2000. **13**(6): p. 413-9.
18. Pereyra, F. and R.H. Rubin, *Prevention and treatment of cytomegalovirus infection in solid organ transplant recipients.* Curr Opin Infect Dis, 2004. **17**(4): p. 357-61.
19. Manicklal, S., et al., *The "silent" global burden of congenital cytomegalovirus.* Clin Microbiol Rev, 2013. **26**(1): p. 86-102.

20. Goderis, J., et al., *Hearing loss and congenital CMV infection: a systematic review*. Pediatrics, 2014. **134**(5): p. 972-82.
21. Ludwig, A. and H. Hengel, *Epidemiological impact and disease burden of congenital cytomegalovirus infection in Europe.* Euro Surveill, 2009. **14**(9): p. 26-32.
22. Hamele, M., et al., *Severe morbidity and mortality with breast milk associated cytomegalovirus infection.* Pediatr Infect Dis J, 2010. **29**(1): p. 84-6.
23. Price, S.M., et al., *Educating women about congenital cytomegalovirus: assessment of health education materials through a web-based survey.* BMC Women's Health, 2014. **14**(1): p. 1-10.
24. Jeon, J., et al., *Knowledge and awareness of congenital cytomegalovirus among women*. Infect Dis Obstet Gynecol, 2006. **2006**.
25. Ross, D.S., et al., *Women's knowledge of congenital cytomegalovirus: results from the 2005 HealthStyles survey*. J Womens Health (Larchmt), 2008. **17**.
26. Jain, M., S. Duggal, and T.D. Chugh, *Cytomegalovirus infection in non-immunosuppressed critically ill patients.* J Infect Dev Ctries, 2011. **5**(8): p. 571-9.
27. Pawelec, G., et al., *The impact of CMV infection on survival in older humans.* Curr Opin Immunol, 2012. **24**(4): p. 507-11.
28. King, A.M.Q., et al., eds. *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. 2012, Elsevier Academic Press.
29. Roizmann, B., et al., *The family Herpesviridae: an update. The Herpesvirus Study Group of the International Committee on Taxonomy of Viruses.* Arch Virol, 1992. **123**(3-4): p. 425-49.
30. Hudson, J.B., *The murine cytomegalovirus as a model for the study of viral pathogenesis and persistent infections.* Arch Virol, 1979. **62**(1): p. 1-29.
31. Reeves, M. and J. Sinclair, *Aspects of human cytomegalovirus latency and reactivation.* Curr Top Microbiol Immunol, 2008. **325**: p. 297-313.
32. Sinclair, J., *Human cytomegalovirus: Latency and reactivation in the myeloid lineage*. J Clin Virol, 2008. **41**(3): p. 180-5.
33. Sinzger, C. and G. Jahn, *Human cytomegalovirus cell tropism and pathogenesis*. Intervirology, 1996. **39**(5-6): p. 302-19.
34. Plachter, B., C. Sinzger, and G. Jahn, *Cell types involved in replication and distribution of human cytomegalovirus*. Adv Virus Res, 1996. **46**: p. 195-261.
35. Mocarski, E.S., et al., *Cytomegaloviruses* in *Fields Virology*, D.M. Knipe and P.M. Howley, Editors. 2013, Lipincott Williams and Wilkins: Philadelphia, PA.
36. Gibson, W., *Structure and formation of the cytomegalovirus virion.* Curr Top Microbiol Immunol, 2008. **325**: p. 187-204.
37. Chen, D.H., et al., *Three-dimensional visualization of tegument/capsid interactions in the intact human cytomegalovirus.* Virology, 1999. **260**(1): p. 10-6.
38. Kalejta, R.F., *Tegument proteins of human cytomegalovirus.* Microbiol Mol Biol Rev, 2008. **72**(2): p. 249-65, table of contents.
39. Stern-Ginossar, N., et al., *Decoding human cytomegalovirus.* Science, 2012. **338**(6110): p. 1088-93.
40. Stern-Ginossar, N., *Decoding viral infection by ribosome profiling*. J Virol, 2015. **89**(12): p. 6164-6.
41. Murphy, E. and T. Shenk, *Human cytomegalovirus genome*. Curr Top Microbiol Immunol, 2008. **325**: p. 1-19.

42. Brocchieri, L., et al., *Predicting coding potential from genome sequence: application to betaherpesviruses infecting rats and mice.* J Virol, 2005. **79**(12): p. 7570-96.
43. Compton, T., D.M. Nowlin, and N.R. Cooper, *Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate.* Virology, 1993. **193**(2): p. 834-41.
44. Isaacson, M.K. and T. Compton, *Human cytomegalovirus glycoprotein B is required for virus entry and cell-to-cell spread but not for virion attachment, assembly, or egress.* J Virol, 2009. **83**(8): p. 3891-903.
45. Vanarsdall, A.L. and D.C. Johnson, *Human cytomegalovirus entry into cells.* Curr Opin Virol, 2012. **2**(1): p. 37-42.
46. Wille, P.T., et al., *Human cytomegalovirus (HCMV) glycoprotein gB promotes virus entry in trans acting as the viral fusion protein rather than as a receptor-binding protein.* MBio, 2013. **4**(3): p. e00332-13.
47. Zhou, M., J.M. Lanchy, and B.J. Ryckman, *Human Cytomegalovirus gH/gL/gO Promotes the Fusion Step of Entry into All Cell Types, whereas gH/gL/UL128-131 Broadens Virus Tropism through a Distinct Mechanism.* J Virol, 2015. **89**(17): p. 8999-9009.
48. Feire, A.L., H. Koss, and T. Compton, *Cellular integrins function as entry receptors for human cytomegalovirus via a highly conserved disintegrin-like domain.* Proc Natl Acad Sci U S A, 2004. **101**(43): p. 15470-5.
49. Feire, A.L., et al., *The glycoprotein B disintegrin-like domain binds beta 1 integrin to mediate cytomegalovirus entry.* J Virol, 2010. **84**(19): p. 10026-37.
50. Isaacson, M.K., A.L. Feire, and T. Compton, *Epidermal growth factor receptor is not required for human cytomegalovirus entry or signaling.* J Virol, 2007. **81**(12): p. 6241-7.
51. Wang, X., et al., *Epidermal growth factor receptor is a cellular receptor for human cytomegalovirus.* Nature, 2003. **424**(6947): p. 456-61.
52. Chan, G., M.T. Nogalski, and A.D. Yurochko, *Activation of EGFR on monocytes is required for human cytomegalovirus entry and mediates cellular motility.* Proc Natl Acad Sci U S A, 2009. **106**(52): p. 22369-74.
53. Ryckman, B.J., et al., *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): p. 710-22.
54. Compton, T., R.R. Nepomuceno, and D.M. Nowlin, *Human cytomegalovirus penetrates host cells by pH-independent fusion at the cell surface.* Virology, 1992. **191**(1): p. 387-95.
55. Ogawa-Goto, K., et al., *Microtubule network facilitates nuclear targeting of human cytomegalovirus capsid.* J Virol, 2003. **77**(15): p. 8541-7.
56. Korioth, F., et al., *The nuclear domain 10 (ND10) is disrupted by the human cytomegalovirus gene product IE1.* Exp Cell Res, 1996. **229**(1): p. 155-8.
57. Maul, G.G., *Initiation of cytomegalovirus infection at ND10.* Curr Top Microbiol Immunol, 2008. **325**: p. 117-32.
58. Wathen, M.W. and M.F. Stinski, *Temporal patterns of human cytomegalovirus transcription: mapping the viral RNAs synthesized at immediate early, early, and late times after infection.* J Virol, 1982. **41**(2): p. 462-77.
59. Pari, G.S., *Nuts and bolts of human cytomegalovirus lytic DNA replication.* Curr Top Microbiol Immunol, 2008. **325**: p. 153-66.

60. Isomura, H. and M.F. Stinski, *Coordination of late gene transcription of human cytomegalovirus with viral DNA synthesis: recombinant viruses as potential therapeutic vaccine candidates.* Expert Opin Ther Targets, 2013. **17**(2): p. 157-66.
61. Weekes, M.P., et al., *Quantitative temporal viromics: an approach to investigate host-pathogen interaction.* Cell, 2014. **157**(6): p. 1460-72.
62. Mettenleiter, T.C., et al., *The way out: what we know and do not know about herpesvirus nuclear egress.* Cell Microbiol, 2013. **15**(2): p. 170-8.
63. Sanchez, V., et al., *Accumulation of virion tegument and envelope proteins in a stable cytoplasmic compartment during human cytomegalovirus replication: characterization of a potential site of virus assembly.* J Virol, 2000. **74**(2): p. 975-86.
64. Mettenleiter, T.C., B.G. Klupp, and H. Granzow, *Herpesvirus assembly: a tale of two membranes.* Curr Opin Microbiol, 2006. **9**(4): p. 423-9.
65. Mettenleiter, T.C., B.G. Klupp, and H. Granzow, *Herpesvirus assembly: an update.* Virus Res, 2009. **143**(2): p. 222-34.
66. Buhler, B., et al., *Characterization of the murine cytomegalovirus early transcription unit e1 that is induced by immediate-early proteins.* J Virol, 1990. **64**(5): p. 1907-19.
67. Wright, D.A. and D.H. Spector, *Posttranscriptional regulation of a class of human cytomegalovirus phosphoproteins encoded by an early transcription unit.* J Virol, 1989. **63**(7): p. 3117-27.
68. Ciocco-Schmitt, G.M., et al., *Identification and characterization of novel murine cytomegalovirus M112-113 (e1) gene products.* Virology, 2002. **294**(1): p. 199-208.
69. Iwayama, S., et al., *Intracellular localization and DNA-binding activity of a class of viral early phosphoproteins in human fibroblasts infected with human cytomegalovirus (Towne strain).* J Gen Virol, 1994. **75 (Pt 12)**: p. 3309-18.
70. Yamamoto, T., et al., *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 Res, 1998. **56**(1): p. 107-14.
71. Penfold, M.E. and E.S. Mocarski, *Formation of cytomegalovirus DNA replication compartments defined by localization of viral proteins and DNA synthesis.* Virology, 1997. **239**(1): p. 46-61.
72. Ahn, J.H., W.J. Jang, and G.S. Hayward, *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).* J Virol, 1999. **73**(12): p. 10458-71.
73. Tang, Q. and G.G. Maul, *Mouse cytomegalovirus immediate-early protein 1 binds with host cell repressors to relieve suppressive effects on viral transcription and replication during lytic infection.* J Virol, 2003. **77**(2): p. 1357-67.
74. Dunn, W., et al., *Functional profiling of a human cytomegalovirus genome.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14223-8.
75. Yu, D., M.C. Silva, and T. Shenk, *Functional map of human cytomegalovirus AD169 defined by global mutational analysis.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 12396-401.
76. Schumacher, U., et al., *Mutations in the M112/M113-coding region facilitate murine cytomegalovirus replication in human cells.* J Virol, 2010. **84**(16): p. 7994-8006.

77. Pari, G.S. and D.G. Anders, *Eleven loci encoding trans-acting factors are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA replication*. J Virol, 1993. **67**(12): p. 6979-88.
78. Iskenderian, A.C., et al., *Four of eleven loci required for transient complementation of human cytomegalovirus DNA replication cooperate to activate expression of replication genes*. J Virol, 1996. **70**(1): p. 383-92.
79. Kagele, D., et al., *Analysis of the interactions of viral and cellular factors with human cytomegalovirus lytic origin of replication, oriLyt*. Virology, 2012. **424**(2): p. 106-14.
80. Kerry, J.A., et al., *The role of ATF in regulating the human cytomegalovirus DNA polymerase (UL54) promoter during viral infection*. J Virol, 1997. **71**(3): p. 2120-6.
81. Li, J., et al., *Major product pp43 of human cytomegalovirus U(L)112-113 gene is a transcriptional coactivator with two functionally distinct domains*. Virology, 1999. **260**(1): p. 89-97.
82. Kim, Y.-E., et al., *Requirement of the N-terminal residues of human cytomegalovirus UL112-113 proteins for viral growth and oriLyt-dependent DNA replication*. Journal of Microbiology, 2015. **53**(8): p. 561-569.
83. Wells, R., L. Stensland, and J. Vieira, *The human cytomegalovirus UL112-113 locus can activate the full Kaposi's sarcoma-associated herpesvirus lytic replication cycle*. J Virol, 2009. **83**(9): p. 4695-9.
84. Park, M.Y., et al., *Interactions among four proteins encoded by the human cytomegalovirus UL112-113 region regulate their intranuclear targeting and the recruitment of UL44 to prereplication foci*. J Virol, 2006. **80**(6): p. 2718-27.
85. Kim, Y.E. and J.H. Ahn, *Role of the specific interaction of UL112-113 p84 with UL44 DNA polymerase processivity factor in promoting DNA replication of human cytomegalovirus*. J Virol, 2010. **84**(17): p. 8409-21.
86. Perez, K.J., et al., *A short cis-acting motif in the M112-113 promoter region is essential for IE3 to activate M112-113 gene expression and is important for murine cytomegalovirus replication*. J Virol, 2013. **87**(5): p. 2639-47.
87. Tang, Q., L. Li, and G.G. Maul, *Mouse cytomegalovirus early M112/113 proteins control the repressive effect of IE3 on the major immediate-early promoter*. J Virol, 2005. **79**(1): p. 257-63.
88. Meier, J.L. and M.F. Stinski, *Major Immediate-Early Enhance and Its Gene Products*, in *Cytomegaloviruses: From Molecular Pathogenesis to Intervention*, M.J. Reddehase, Editor. 2013, Caister Academic Press: Norfolk, UK.
89. Arai, Y., et al., *Neuron-specific activation of murine cytomegalovirus early gene e1 promoter in transgenic mice*. Am J Pathol, 2003. **163**(2): p. 643-52.
90. Mutnal, M.B., S. Hu, and J.R. Lokensgard, *Persistent humoral immune responses in the CNS limit recovery of reactivated murine cytomegalovirus*. PLoS One, 2012. **7**(3): p. e33143.
91. Crick, F., *Split genes and RNA splicing*. Science, 1979. **204**(4390): p. 264-71.
92. Sharp, P.A., *The discovery of split genes and RNA splicing*. Trends Biochem Sci, 2005. **30**(6): p. 279-81.
93. Black, D.L., *Mechanisms of alternative pre-messenger RNA splicing*. Annu Rev Biochem, 2003. **72**: p. 291-336.
94. Patel, A.A. and J.A. Steitz, *Splicing double: insights from the second spliceosome*. Nat Rev Mol Cell Biol, 2003. **4**(12): p. 960-70.
95. Sharp, P.A., *Split genes and RNA splicing*. Cell, 1994. **77**(6): p. 805-15.

96. Konarska, M.M., et al., *Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors*. Nature, 1985. **313**(6003): p. 552-7.
97. Early, P., et al., *Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways*. Cell, 1980. **20**(2): p. 313-9.
98. Matlin, A.J., F. Clark, and C.W. Smith, *Understanding alternative splicing: towards a cellular code*. Nat Rev Mol Cell Biol, 2005. **6**(5): p. 386-98.
99. Kornblihtt, A.R., et al., *Alternative splicing: a pivotal step between eukaryotic transcription and translation*. Nat Rev Mol Cell Biol, 2013. **14**(3): p. 153-65.
100. David, C.J. and J.L. Manley, *The search for alternative splicing regulators: new approaches offer a path to a splicing code*. Genes Dev, 2008. **22**(3): p. 279-85.
101. Srebrow, A. and A.R. Kornblihtt, *The connection between splicing and cancer*. Journal of Cell Science, 2006. **119**(13): p. 2635-2641.
102. Sammeth, M., S. Foissac, and R. Guigo, *A general definition and nomenclature for alternative splicing events*. PLoS Comput Biol, 2008. **4**(8): p. e1000147.
103. Ward, A.J. and T.A. Cooper, *The pathobiology of splicing*. J Pathol, 2010. **220**(2): p. 152-63.
104. Sandri-Goldin, R.M., *Viral regulation of mRNA export*. J Virol, 2004. **78**(9): p. 4389-96.
105. Gatherer, D., et al., *High-resolution human cytomegalovirus transcriptome*. Proc Natl Acad Sci U S A, 2011. **108**(49): p. 19755-60.
106. Ma, Y., et al., *Human CMV transcripts: an overview*. Future Microbiol, 2012. **7**(5): p. 577-93.
107. Stenberg, R.M., P.R. Witte, and M.F. Stinski, *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, 1985. **56**(3): p. 665-75.
108. Stamminger, T., E. Puchtler, and B. Fleckenstein, *Discordant expression of the immediate-early 1 and 2 gene regions of human cytomegalovirus at early times after infection involves posttranscriptional processing events*. J Virol, 1991. **65**(5): p. 2273-82.
109. Rawlinson, W.D. and B.G. Barrell, *Spliced transcripts of human cytomegalovirus*. J Virol, 1993. **67**(9): p. 5502-13.
110. Kerry, J.A., et al., *Isolation and characterization of a low-abundance splice variant from the human cytomegalovirus major immediate-early gene region*. J Virol, 1995. **69**(6): p. 3868-72.
111. Awasthi, S., J.A. Isler, and J.C. Alwine, *Analysis of splice variants of the immediate-early 1 region of human cytomegalovirus*. J Virol, 2004. **78**(15): p. 8191-200.
112. Tenney, D.J. and A.M. Colberg-Poley, *Expression of the human cytomegalovirus UL36-38 immediate early region during permissive infection*. Virology, 1991. **182**(1): p. 199-210.
113. Colberg-Poley, A.M., *Functional roles of immediate early proteins encoded by the human cytomegalovirus UL36-38, UL115-119, TRS1/IRS1 and US3 loci*. Intervirology, 1996. **39**(5-6): p. 350-60.
114. Adair, R., G.W. Liebisch, and A.M. Colberg-Poley, *Complex alternative processing of human cytomegalovirus UL37 pre-mRNA*. J Gen Virol, 2003. **84**(Pt 12): p. 3353-8.

115. Adair, R., et al., *Alteration of cellular RNA splicing and polyadenylation machineries during productive human cytomegalovirus infection.* J Gen Virol, 2004. **85**(Pt 12): p. 3541-53.
116. McCormick, A.L., *Control of apoptosis by human cytomegalovirus.* Curr Top Microbiol Immunol, 2008. **325**: p. 281-95.
117. Liu, W., Y. Zhao, and B. Biegalke, *Analysis of human cytomegalovirus US3 gene products.* Virology, 2002. **301**(1): p. 32-42.
118. Liu, Z., M. Winkler, and B. Biegalke, *Human cytomegalovirus: host immune modulation by the viral US3 gene.* Int J Biochem Cell Biol, 2009. **41**(3): p. 503-6.
119. Noriega, V.M., et al., *Human cytomegalovirus US3 modulates destruction of MHC class I molecules.* Mol Immunol, 2012. **51**(2): p. 245-53.
120. Mach, M., et al., *Complex formation by glycoproteins M and N of human cytomegalovirus: structural and functional aspects.* J Virol, 2005. **79**(4): p. 2160-70.
121. Juranic Lisnic, V., et al., *Dual analysis of the murine cytomegalovirus and host cell transcriptomes reveal new aspects of the virus-host cell interface.* PLoS Pathog, 2013. **9**(9): p. e1003611.
122. Lin, Z., et al., *Quantitative and qualitative RNA-Seq-based evaluation of Epstein-Barr virus transcription in type I latency Burkitt's lymphoma cells.* J Virol, 2010. **84**(24): p. 13053-8.
123. van Beurden, S.J., et al., *Anguillid herpesvirus 1 transcriptome.* J Virol, 2012. **86**(18): p. 10150-61.
124. Arias, C., et al., *KSHV 2.0: a comprehensive annotation of the Kaposi's sarcoma-associated herpesvirus genome using next-generation sequencing reveals novel genomic and functional features.* PLoS Pathog, 2014. **10**(1): p. e1003847.
125. Olah, P., et al., *Characterization of pseudorabies virus transcriptome by Illumina sequencing.* BMC Microbiol, 2015. **15**: p. 130.
126. Kerry, J.A., et al., *Multiple regulatory events influence human cytomegalovirus DNA polymerase (UL54) expression during viral infection.* J Virol, 1996. **70**(1): p. 373-82.
127. Freund, M., et al., *A novel approach to describe a U1 snRNA binding site.* Nucleic Acids Res, 2003. **31**(23): p. 6963-75.
128. Tischer, B.K., G.A. Smith, and N. Osterrieder, *En passant mutagenesis: a two step markerless red recombination system.* Methods Mol Biol, 2010. **634**: p. 421-30.
129. Valchanova, R.S., et al., *Murine cytomegalovirus m142 and m143 are both required to block protein kinase R-mediated shutdown of protein synthesis.* J Virol, 2006. **80**(20): p. 10181-90.
130. Mack, C., et al., *Inhibition of proinflammatory and innate immune signaling pathways by a cytomegalovirus RIP1-interacting protein.* Proc Natl Acad Sci U S A, 2008. **105**(8): p. 3094-9.
131. Budt, M., et al., *Specific inhibition of the PKR-mediated antiviral response by the murine cytomegalovirus proteins m142 and m143.* J Virol, 2009. **83**(3): p. 1260-70.
132. Mellacheruvu, D., et al., *The CRAPome: a contaminant repository for affinity purification-mass spectrometry data.* Nat Methods, 2013. **10**(8): p. 730-6.
133. Wright, D.A., S.I. Staprans, and D.H. Spector, *Four phosphoproteins with common amino termini are encoded by human cytomegalovirus AD169.* J Virol, 1988. **62**(1): p. 331-40.

134. Wang, S.K., et al., *Calpains mediate the proteolytic modification of human cytomegalovirus UL112-113 proteins*. J Gen Virol, 2015. **96**(Pt 5): p. 1115-26.
135. Kohler, A. and E. Hurt, *Exporting RNA from the nucleus to the cytoplasm*. Nat Rev Mol Cell Biol, 2007. **8**(10): p. 761-73.
136. Valencia, P., A.P. Dias, and R. Reed, *Splicing promotes rapid and efficient mRNA export in mammalian cells*. Proc Natl Acad Sci U S A, 2008. **105**(9): p. 3386-91.
137. Johnson, L.A. and R.M. Sandri-Goldin, *Efficient nuclear export of herpes simplex virus 1 transcripts requires both RNA binding by ICP27 and ICP27 interaction with TAP/NXF1*. J Virol, 2009. **83**(3): p. 1184-92.
138. Sandri-Goldin, R.M., *Nuclear export of herpes virus RNA*. Curr Top Microbiol Immunol, 2001. **259**: p. 2-23.
139. Farjot, G., et al., *Epstein-Barr virus EB2 protein exports unspliced RNA via a Crm-1-independent pathway*. J Virol, 2000. **74**(13): p. 6068-76.
140. Hiriart, E., et al., *A novel nuclear export signal and a REF interaction domain both promote mRNA export by the Epstein-Barr virus EB2 protein*. J Biol Chem, 2003. **278**(1): p. 335-42.
141. Toth, Z. and T. Stamminger, *The human cytomegalovirus regulatory protein UL69 and its effect on mRNA export*. Front Biosci, 2008. **13**: p. 2939-49.
142. Lischka, P., et al., *The UL69 transactivator protein of human cytomegalovirus interacts with DEXD/H-Box RNA helicase UAP56 to promote cytoplasmic accumulation of unspliced RNA*. Mol Cell Biol, 2006. **26**(5): p. 1631-43.
143. Zielke, B., et al., *Transfer of the UAP56 interaction motif of human cytomegalovirus pUL69 to its murine cytomegalovirus homolog converts the protein into a functional mRNA export factor that can substitute for pUL69 during viral infection*. J Virol, 2012. **86**(13): p. 7448-53.
144. Domingo, E., J. Sheldon, and C. Perales, *Viral quasispecies evolution*. Microbiol Mol Biol Rev, 2012. **76**(2): p. 159-216.
145. Ghazal, P., et al., *Elimination of ie1 significantly attenuates murine cytomegalovirus virulence but does not alter replicative capacity in cell culture*. J Virol, 2005. **79**(11): p. 7182-94.
146. Stahl, S., et al., *Cytomegalovirus downregulates IRE1 to repress the unfolded protein response*. PLoS Pathog, 2013. **9**(8): p. e1003544.
147. Loh, L.C., et al., *Sequence analysis and expression of the murine cytomegalovirus phosphoprotein pp50, a homolog of the human cytomegalovirus UL44 gene product*. Virology, 1994. **200**(2): p. 413-27.
148. Martinez, F.P., R.S. Cosme, and Q. Tang, *Murine cytomegalovirus major immediate-early protein 3 interacts with cellular and viral proteins in viral DNA replication compartments and is important for early gene activation*. J Gen Virol, 2010. **91**(Pt 11): p. 2664-76.
149. Siomi, H. and G. Dreyfuss, *A nuclear localization domain in the hnRNP A1 protein*. J Cell Biol, 1995. **129**(3): p. 551-60.
150. Papadopoulou, C., et al., *Expression profile and interactions of hnRNP A3 within hnRNP/mRNP complexes in mammals*. Arch Biochem Biophys, 2012. **523**(2): p. 151-60.
151. Adamson, B., et al., *A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response*. Nat Cell Biol, 2012. **14**(3): p. 318-28.
152. Dreyfuss, G., et al., *hnRNP proteins and the biogenesis of mRNA*. Annu Rev Biochem, 1993. **62**: p. 289-321.

153. Martinez-Arribas, F., et al., *Positive correlation between the expression of X-chromosome RBM genes (RBMX, RBM3, RBM10) and the proapoptotic Bax gene in human breast cancer*. J Cell Biochem, 2006. **97**(6): p. 1275-82.
154. Jean-Philippe, J., S. Paz, and M. Caputi, *hnRNP A1: the Swiss army knife of gene expression*. Int J Mol Sci, 2013. **14**(9): p. 18999-9024.
155. Tanaka, E., et al., *HnRNP A3 binds to and protects mammalian telomeric repeats in vitro*. Biochem Biophys Res Commun, 2007. **358**(2): p. 608-14.
156. Roux, K.J., D.I. Kim, and B. Burke, *BioID: a screen for protein-protein interactions*. Curr Protoc Protein Sci, 2013. **74**: p. Unit 19 23.
157. 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): p. 359-68.
158. Alvira, M.R., et al., *Genetic studies exposing the splicing events involved in herpes simplex virus type 1 latency-associated transcript production during lytic and latent infection*. J Virol, 1999. **73**(5): p. 3866-76.
159. Kulesza, C.A. and T. Shenk, *Human cytomegalovirus 5-kilobase immediate-early RNA is a stable intron*. J Virol, 2004. **78**(23): p. 13182-9.
160. Kulesza, C.A. and T. Shenk, *Murine cytomegalovirus encodes a stable intron that facilitates persistent replication in the mouse*. Proc Natl Acad Sci U S A, 2006. **103**(48): p. 18302-7.
161. Harvey, D.M. and A.J. Levine, *p53 alteration is a common event in the spontaneous immortalization of primary BALB/c murine embryo fibroblasts*. Genes Dev, 1991. **5**(12B): p. 2375-85.
162. Swift, S., et al., *Rapid production of retroviruses for efficient gene delivery to mammalian cells using 293T cell-based systems*. Curr Protoc Immunol, 2001. **Chapter 10**: p. Unit 10 17C.
163. O'Connor, C.M. and E.A. Murphy, *A myeloid progenitor cell line capable of supporting human cytomegalovirus latency and reactivation, resulting in infectious progeny*. J Virol, 2012. **86**(18): p. 9854-65.
164. Redwood, A.J., et al., *Use of a murine cytomegalovirus K181-derived bacterial artificial chromosome as a vaccine vector for immunocontraception*. J Virol, 2005. **79**(5): p. 2998-3008.
165. Krause, E., et al., *Murine cytomegalovirus virion-associated protein M45 mediates rapid NF-kappaB activation after infection*. J Virol, 2014. **88**(17): p. 9963-75.
166. Angulo, A., P. Ghazal, and M. Messerle, *The major immediate-early gene ie3 of mouse cytomegalovirus is essential for viral growth*. J Virol, 2000. **74**(23): p. 11129-36.
167. Kalejta, R.F., J.T. Bechtel, and T. Shenk, *Human cytomegalovirus pp71 stimulates cell cycle progression by inducing the proteasome-dependent degradation of the retinoblastoma family of tumor suppressors*. Mol Cell Biol, 2003. **23**(6): p. 1885-95.
168. Jurak, I. and W. Brune, *Induction of apoptosis limits cytomegalovirus cross-species infection*. EMBO J, 2006. **25**(11): p. 2634-42.
169. Birnboim, H.C. and J. Doly, *A rapid alkaline extraction procedure for screening recombinant plasmid DNA*. Nucleic Acids Res, 1979. **7**(6): p. 1513-23.
170. Osborn, J.E. and D.L. Walker, *Enhancement of infectivity of murine cytomegalovirus in vitro by centrifugal inoculation*. J Virol, 1968. **2**(9): p. 853-8.
171. Flint, S.J., et al., *Principles of Virology*. 2009: ASM Press.

172. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.
173. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.
174. Hoffmann, S., et al., *Fast mapping of short sequences with mismatches, insertions and deletions using index structures.* PLoS Comput Biol, 2009. **5**(9): p. e1000502.

10 Appendix

10.1 Publications and presentations

10.1.1 Publications

Parts of this thesis were published in:

Functional dissection of an alternatively spliced herpesvirus gene by splice site mutagenesis
Schommartz T, Loroch S, Alawi M, Grundhoff A, Sickmann A, Brune W.
J Virol. 2016 Apr 14;90(9):4626-36.

The author of this thesis contributed to the following publications during the course of this study.

Functional Comparison of the Molluscum Contagiosum Virus MC159/vFLIP with the Murine Cytomegalovirus M36/vICA and M45/vIRA Proteins.
Hüttmann J, Krause E, Schommartz T, Brune W.
J Virol. 2015 Dec 30;90(6):2895-905.

Calcium spirulan derived from Spirulina platensis inhibits herpes simplex virus 1 attachment to human keratinocytes and protects against herpes labialis.
Mader J, Gallo A, Schommartz T, Handke W, Nagel CH, Günther P, Brune W, Reich K.
J Allergy Clin Immunol. 2016 Jan; 137(1):197-203.

Multidimensional electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) for quantitative analysis of the proteome and phosphoproteome in clinical and biomedical research.
Loroch S, Schommartz T, Brune W, Zahedi RP, Sickmann A.
Biochim Biophys Acta. 2015 May;1854(5):460-8.

The author of this thesis contributed to the following publications before the start of this study.

T cell-independent B cell activation induces immunosuppressive sialylated IgG antibodies.
Hess C, Winkler A, Lorenz AK, Holecska V, Blanchard V, Eiglmeier S, Schoen AL, Bitterling J, Stoehr AD, Petzold D, Schommartz T, Mertes MM, Schoen CT, Tiburzy B, Herrmann A, Köhl J, Manz RA, Madaio MP, Berger M, Wardemann H, Ehlers M.
J Clin Invest. 2013 Sep;123(9):3788-96.

Tolerance induction with T cell-dependent protein antigens induces regulatory sialylated IgGs.
Oefner CM, Winkler A, Hess C, Lorenz AK, Holecska V, Huxdorf M, Schommartz T, Petzold D, Bitterling J, Schoen AL, Stoehr AD, Vu Van D, Darcan-Nikolaisen Y, Blanchard V, Schmudde I, Laumonnier Y, Ströver HA, Hegazy AN, Eiglmeier S, Schoen CT, Mertes MM, Loddenkemper C, Löhning M, König P, Petersen A, Luger EO, Collin M, Köhl J, Hutloff A, Hamelmann E, Berger M, Wardemann H, Ehlers M.
J Allergy Clin Immunol. 2012 Jun;129(6):1647-55.

TLR9 in peritoneal B-1b cells is essential for production of protective self-reactive IgM to control Th17 cells and severe autoimmunity.
Stoehr AD, Schoen CT, Mertes MM, Eiglmeier S, Holecska V, Lorenz AK, Schommartz T, Schoen AL, Hess C, Winkler A, Wardemann H, Ehlers M.
J Immunol. 2011 Sep 15;187(6):2953-65.

10.1.2 Presentations

The author presented parts of this study at the following conferences:

March 2013	23rd Annual meeting of the Society of Virology, Kiel, Germany	(poster)
October 2014	HPI Joint Scientific Retreat Hamburg, Germany	(oral presentation)
November 2014	2nd Waldthausen Castle Symposium, Mainz, Germany	(poster)
March 2015	24th Annual Meeting of the Society of Virology, Bochum, Germany	(poster)
May 2015	3rd International Symposium of the Virtual Institute VISTRIE Braunschweig, Germany	(poster)
July 2015	40th Annual International Herpesvirus Workshop, Boise, Idaho, USA (registration awardee)	(oral presentation)
October 2015	HPI Joint Scientific Retreat Hamburg, Germany	(oral presentation)

10.2 Curriculum vitae

Entfällt aus Datenschutzgründen.

10.3 List of abbreviations

AIDS	acquired immunodeficiency syndrome
bp	base pair(s)
CD	cluster of differentiation
CMV	cytomegalovirus
CNS	central nervous system
CPE	cytopathic effect
dsDNA	double stranded DNA
E	early
EBV	Epstein-Barr virus
E-L	early-late
ER	endoplasmic reticulum
GFP	green fluorescent protein
HAART	highly active antiretroviral therapy
HCMV	human cytomegalovirus
HIV	human immunodeficiency virus
hnRNP	heterogeneous nuclear ribonucleoprotein
hpi	hours post infection
HPLC	high performance liquid chromatography
HSV	herpes simplex virus
IB	immunoblot
IE	immediate early
IE3AM	immediate early 3 activating motif
IP	immunoprecipitation
IRS	internal repeat short
kb	kilo base pairs (1000 bp)
KSHV	Kaposi's sarcoma-associated herpesvirus
L	late
LC-MS	liquid chromatography–mass spectrometry
MCMV	murine cytomegalovirus
MIEP	major immediate early promotor
MOI	multiplicity of infection
MS	mass spectrometry
ND10	nuclear domain 10

NLS	nuclear localization signal
*ori*Lyt	origin of replication (lytic)
PCR	polymerase chain reaction
PML	promyelocytic leukemia (associated)
PSM	peptide spectrum match
RBMX	RNA Binding Motif Protein, X-Linked
Rev	revertant
RNAi	RNA interference
SILAC	stable isotope labeling with amino acids in cell culture
snRNP	small nuclear ribonucleo protein
$TCID_{50}$	tissue culture infectious dose 50 %
TEV	tobacco etch virus
TRS	terminal repeat short
UL	unique long
US	unique short
WB	western blot
wt	wild type

10.4 Toxicity of chemicals

substance	GHS symbol	hazard statements	precautionary statements
2-mercaptoethanol		H301 + H331-H310-H315-H317-H318-H373-H410	P261-P280-P301 + P310 + P330-P302 + P352 + P310-P305 + P351 + P338 + P310-P403 + P233
acetic acid		H226-H314	P280-P305 + P351 + P338-P310
acrylamide		H301-H312 + H332-H315-H317-H319-H340-H350-H361f-H372	P201-P280-P301 + P310-P305 + P351 + P338-P308 + P313
ammonium bicarbonate		H302	P301 + P312 + P330
ammonium persulfate		H272-H302-H315-H317-H319-H334-H335	P220-P261-P280-P305 + P351 + P338-P342 + P311
ampicillin		H315-H317-H319-H334-H335	P261-P280-P305 + P351 + P338-P342 + P311
bis-acrylamide		H302 + H332	
boric acid		H360FD	P201-P308 + P313
chloramphenicol		H350	P201-P308 + P313
deoxycholate		H302-H335	P301 + P312 + P330
diphtheria toxin		H300	P264-P301 + P310
dithiothreitol		H302-H315-H319-H335	P261-P305 + P351 + P338
EDTA		H319	P305 + P351 + P338
ethanol		H225-H319	P210-P280-P305 + P351 + P338-P337 + P313-P403 + P235

ethidium bromide		H302-H330-H341	P260-P281-P284-P310
guanidine hydrochloride		H302 + H332-H315-H319	P261-P280-P301 + P312 + P330-P304 + P340 + P312-P305 + P351 + P338-P337 + P313
hydrochloric acid		H290-H314-H335	P261-P280-P305 + P351 + P338-P310
hygromycin B		H300 +H310 + H330-H318-H334	P260-P264-P280-P284-P301 + P310-P302 + P350
iodoacetamide		H301-H317-H334	P261-P280-P301 + P310-P342 + P311
isopropanol		H225-H319-H336	P210-P261-P305 + P351 + P338
kanamycin		H360	P201-P308 + P313
liquid nitrogen		H281	P202-P271 + P403-P282
methanol		H225-H301 + H311 + H331-H370	P210-P260-P280-P301 + P310-P311
penicillin		H317-H334	P261-P280-P342 + P311
protein A-agarose		H226	
protein G-agarose		H226	
puromycin		H373	
sodium dodecyl sulfate		H315-H318-H335	P280-P304 + P340 + P312-P305 + P351 + P338 + P310
sodium hydroxide		H290-H314	P280-P305 + P351 + P338-P310
streptomycin		H302-H361	P281

TEMED		H225-H302-H314-H332	P210-P280-P305 + P351 + P338-P310
trifluoroacetic acid		H314-H332-H412	P261-P273-P280-P303 + P361 + P353-P304 + P340 + P310-P305 + P351 + P338
Triton X-100		H302-H319-H411	P273-P280-P301 + P312 + P330-P337 + P313-P391-P501

10.5 Acknowledgments

Zuallererst möchte ich mich bei meinem Doktorvater Prof. Dr. Wolfram Brune für die Überlassung dieses spannenden Forschungsthemas bedanken. Er gab mir den Freiraum selbständig zu arbeiten und eigene Ideen zu entwickeln. Trotzdem stand er mir, wann immer es nötig war, mit hilfreichen Ratschlägen zur Seite. Nur dank unzähliger konstruktiver Diskussionen mit ihm, konnte die Arbeit in dieser Form entstehen.

Mein Dank gilt auch Herrn Dr. Markus Perbandt für die Zweitbetreuung meiner Doktorarbeit. Er unterstützte mich durch hilfreiche Fragen und Diskussionen.

Herrn Prof. Dr. Ulrich Hahn möchte ich für die Übernahme des Zweitgutachtens der Dissertation danken. Bei Herrn Prof. Dr. Wolfgang Maison und Frau Dr. Maria Riedner bedanke ich mich für die Begutachtung der Disputation.

I also want to thank Prof. Julie A. Kerry, PhD; Prof. Stipan Jonjić, MD, PhD and Prof. Lambert Loh, for providing antibodies crucial for this study.

Ich danke Frau Dr. Daniela Indenbirken und Herrn Malik Alawi für die exzellente Unterstützung mit den RNASeq Experimenten.

Herrn Stefan Loroch und Herrn Prof. Dr. Albert Sickmann danke ich für die Unterstützung bei den massenspektrometrischen Versuchen. Stefan, dank dir habe ich ein grundlegendes Verständnis für Proteomics entwickelt. Ich habe in dir nicht nur einen großartigen Kooperationspartner, sondern auch einen Freund gewonnen. Danke!

Mein Dank gilt auch meinen beiden fantastischen Masterstudenten Frau Rebekka Brost und Frau Julia Hüttmann die ich betreuen durfte. Es ist das größte Kompliment, wenn der Lehrer am Ende von seinen Schülern lernen kann. Danke für die super Zeit!

I have to send special thanks to Dr. Eva Krause, Dr. Eleonore Ostermann, Leena Ukil, PhD, Theo Potgieter, Dominik Wigger and Akshita Patil for critical reading of this manuscript. All of you helped to drastically improve the quality of this thesis. All remaining mistakes are the fault of my own.

A huge thank you to all present and former lab members of the HPI research unit virus host interaction. Ana, Antonio, Bing, Christina, Doris, Elena, Eleonore, Eva, Florian, Gabi, Jiajia, Julia, Kerstin, Leila, Martina, Olga, Patricia, Renke, Sebastian, Theo and Wiebke you made working in the lab great every day. Thanks for good cooperation and lots of fun inside and outside of the lab!

Ein besonderer Dank gilt meiner gesamten Familie, welche mich auf meinem bisherigen Lebensweg stets und ständig unterstützt hat. Sie standen mir in schwierigen Zeiten immer zur Seite. Ohne euch geht es nicht.

Zuletzt möchte ich mich bei meiner wundervollen Frau Linda bedanken. Worte reichen nicht aus um zu sagen, wie wichtig du bist. Danke, dass du immer an meiner Seite bist!

10.6 Eidesstattliche Versicherung

Hiermit versichere ich, Tim Schommartz geboren am 28.12.1985 in Rostock, an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.

Hamburg, Februar 2016

Tim Schommartz