Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Biologia

SCUOLA DI DOTTORATO DI RICERCA IN BIOSCIENZE E BIOTECNOLOGIE
INDIRIZZO: Biologia Cellulare
CICLO XXVII

Characterization of pUL5, an HCMV protein interacting with the cellular protein IQGAP1

Direttore della Scuola: Ch.mo Prof. Giuseppe Zanotti
Coordinatore d'indirizzo: Ch.mo Prof. Paolo Bernardi
Supervisore: Ch.mo Prof. Cesare Montecucco
Co-supervisore: Dr. Yasushi Uematsu

Dottorando: Giulia Anselmi

TABLE OF CONTENTS

RIASSUNTO IN ITALIANO

Il citomegalovirus umano (HCMV) è un virus che infetta la maggioranza degli adulti nel mondo. In pazienti sani, HCMV stimola una forte risposta immunitaria in grado di limitare e contenere il diffondersi dell'infezione, sebbene questo virus sia capace di restare latente nelle cellule infettate e di riattivarsi spontaneamente quando le condizioni sono più favorevoli. La forma acuta dell'infezione da HCMV è associata a particolari condizioni in cui il sistema immunitario è in qualche modo compromesso, ad esempio in pazienti che hanno ricevuto un trapianto d'organo, in pazienti affetti da sindrome da immunodeficienza acquisita (AIDS) o in feti che contraggono l'infezione congenita. HCMV possiede il più grande genoma tra gli herpesvirus e, nel tempo, la comunità scientifica sta cercando di esplorare il suo enorme potenziale di espressione.

In questo studio ci siamo occupati del prodotto proteico dell'*open reading frame* (ORF) *UL5*, appartenente alla famiglia genica RL11 e ancora da caratterizzare. *UL5* è presente soltanto nei citomegalovirus di umani e scimpanzé mentre è assente nei citomegalovirus che infettano altre specie e viene pertanto considerato un gene 'privato'. La sequenza amminoacidica di *UL5* è abbastanza conservata tra i diversi ceppi del citomegalovirus umano e tutte le differenze, benché minime, sono concentrate all'N-terminale. Predizioni *in silico* classificano pUL5 come una proteina di membrana di tipo Ib mancante di un peptide segnale e senza siti di N-glicosilazione, mentre sono predetti 7 siti di O-glicosilazione. In letteratura sono descritti tre mRNA di lunghezza simile contenenti *UL5* insieme all'ORF precedente *UL4*. Per confermare questo ed escludere la presenza di altri trascritti contenenti *UL5*, abbiamo eseguito un Northern blot sull'RNA totale di cellule infettate da HCMV, estratto a vari tempi dopo l'infezione usando una sonda a DNA specifica per *UL5*. Il saggio ha evidenziato un'unica banda di circa 1.5 kb, escludendo la presenza di altri trascritti contenenti *UL5*. Allo scopo di studiare il prodotto proteico di *UL5*, abbiamo generato un virus ricombinante in cui una serie di tag ottimizzata (2StrepII-2FLAG) è stata aggiunta al C-terminale di pUL5 tramite

la mutagenesi a due step di un cromosoma batterico artificiale (BAC) contenente l'intero genoma del ceppo TR di HCMV. Per studiare il profilo di espressione di pUL5 è stato eseguito un immunoblot con un anticorpo anti-FLAG su lisati di fibroblasti infettati con il virus ricombinante, preparati a diversi tempi dopo l'infezione. Il Western blot ha rilevato la presenza di due specie di circa 24 e 15 kDa, espresse con cinetica *late* e *early*, rispettivamente. Poiché il peso molecolare atteso della proteina intera compresa di tag è 24.5 kDa, è probabile che questa corrisponda alla banda a peso più alto, suggerendo che non esistano forme glicosilate di pUL5. Per confermare questa teoria abbiamo effettuato un'analisi di deglicosilazione sia in trasfezione che in infezione e non abbiamo rilevato alcuna differenza nell'altezza delle bande corrispondenti a pUL5. Per indagare sulla natura della banda a più basso peso molecolare abbiamo mutato un secondo ATG presente nella sequenza codificante di *UL5* che potrebbe corrispondere ad un secondo codone d'inizio della traduzione. Abbiamo poi trasfettato delle cellule HEK293T con il plasmide di *UL5* mutata ed eseguito un Western blot sui lisati, rilevando una forte riduzione nel segnale della banda a 15 kDa. Inoltre, abbiamo prodotto in *E. coli* una forma ricombinante di pUL5 mancante del tratto trans-membrana, l'abbiamo purificata e utilizzata per produrre un antisiero nei topi. Ripetendo gli esperimenti utilizzando il virus *wild-type* e l'antisiero di topo, abbiamo ottenuto gli stessi risultati, escludendo la possibilità di artefatti dovuti alla presenza dei tag. Esperimenti di microscopia confocale su fibroblasti infettati sia con il virus *wild-type* che con quello ricombinante e utilizzando sia l'antisiero di topo che l'anticorpo anti-FLAG hanno rilevato pUL5 in co-localizzazione con marcatori del trans-Golgi network e delle vie secretorie, in un compartimento formato tipicamente durante l'infezione da HCMV e deputato all'assemblaggio dei nuovi virus. Allo scopo di verificare se pUL5 viene incorporata nei virioni maturi, abbiamo effettuato un Western blot su virioni purificati e non abbiamo riscontrato la presenza della proteina, che probabilmente viene espressa soltanto durante il ciclo di infezione ma non costituisce una componente strutturale del virus. Per cercare di capire quale possa essere il ruolo di pUL5 durante l'infezione, abbiamo creato un virus

ricombinante in cui l'intera ORF di *UL5* è stata deleta mediante mutagenesi del BAC e abbiamo valutato la sua capacità di replicazione in fibroblasti umani. Il virus mutante *knock-out* di *UL5* si è rivelato ugualmente in grado di infettare fibroblasti ma con una cinetica di replicazione fortemente ridotta. Infine, abbiamo effettuato un esperimento di *pull-down* di pUL5 con il fine di identificare eventuali interattori. A questo scopo, abbiamo infettato dei fibroblasti umani con il virus ricombinante taggato e abbiamo immunoprecipitato pUL5 utilizzando delle palline magnetiche rivestite di anticorpo anti-FLAG. L'analisi per spettrometria di massa dell'eluato ha portato all'identificazione della proteina cellulare IQGAP1 come uno degli interattori di pUL5. IQGAP1 è una proteina cellulare espressa in maniera ubiquitaria e coinvolta in molti meccanismi cellulari, tra cui il rimodellamento del citoscheletro, l'adesione cellulare e la regolazione del ciclo cellulare, oltre ad essere un noto bersaglio per i patogeni intracellulari. Questa interazione è stata ulteriormente confermata da immunoprecipitazioni reciproche in trasfezione, dando un ulteriore indizio su quale potrebbe essere il ruolo di pUL5 durante il ciclo di infezione di HCMV.

SUMMARY

Human cytomegalovirus (HCMV) is a ubiquitous virus infecting the majority of adults worldwide. In healthy individuals, a strong immune response to HCMV is able to limit and contain the spread of the disease, although the virus is able to establish a lifelong lasting latency with recurrent and spontaneous reactivation. Acute disease is observed only in particular settings where the normal immune response of the patient is weak or compromised, such as in transplant or patients with Acquired Immuno-Deficiency Syndrome (AIDS) or in fetuses that acquire the infection congenitally. HCMV possesses the largest genome among herpesviruses and a lot of effort is being put on the study of the huge coding potential this virus displays in order to have a better understanding of the mechanisms behind its infection.

In this work, we address the study of the protein product of *UL5* open reading frame (ORF), belonging to the RL11 gene family and previously uncharacterized. *UL5* is only present in human and chimpanzee cytomegaloviruses while is absent in cytomegaloviruses infecting other species, therefore is considered a 'private' gene. Among HCMV strains, *UL5* aminoacid sequence is relatively conserved and all minor differences are concentrated in the N-terminus. *In silico* predictions classify pUL5 as a type Ib membrane protein with no signal peptide and no N-linked glycosylations, while 7 O-glycosylation sites are predicted. Three different mRNAs with similar lengths including *UL5* and the previous ORF *UL4* have already been described in literature. In order to confirm these data and to exclude the presence of other *UL5*-containing transcripts, we performed a Northern blot analysis on the total RNA extracted from HCMV-infected cells at different times post infection (p.i.) using an *UL5*-specific biotinylated DNA probe. A band around 1.5 kb was detected, excluding the presence of further transcripts. To address the study of UL5 protein product (pUL5), a two-step mutagenesis was carried out on a bacterial artificial chromosome (BAC) containing the entire genome of the TR strain of HCMV to generate a recombinant virus with the addition of an optimized combination of tags (2StrepII-2FLAG) at the C-terminus of pUL5. pUL5

expression profile was studied by immunoblot on lysates of fibroblasts infected with the recombinant pUL5-tagged virus prepared at different times p.i. using an anti-FLAG antibody. Two species of approximately 24 and 15 kDa were detected with late and early expression kinetics, respectively. Since the predicted molecular weight of the entire tagged protein is 24.5 kDa, the higher molecular weight form probably corresponds to the full-length protein. Glycosylation analysis of pUL5 both in transfection and in infection further confirmed the absence of glycosylation events. To investigate the nature of the lower molecular weight species, we performed a PCR site-directed mutagenesis to eliminate a second ATG in *UL5* ORF. Western blot analysis on lysates of HEK293T cells transfected with the mutated plasmid of *UL5* revealed a strong reduction in the signal of the 15 kDa species. A recombinant form of pUL5 lacking the putative transmembrane region was expressed in *E. coli*, purified and used to obtain an antiserum in mice. Western blot on lysates of fibroblasts infected with the wild-type virus using the mouse antiserum replicated the previous data, excluding a possible influence of the tags to the obtained results. Confocal microscopy of infected fibroblasts both with wild-type and recombinant viruses, implementing both the pUL5 antiserum and the anti-FLAG antibody, localizes pUL5 with markers of the trans-Golgi network and secretory pathways, in the compartment of viral assembly. Western blot on purified virions failed to reveal the presence of pUL5, suggesting that this protein is only expressed in the host cell during infection but it is not incorporated in the released viral progeny. To have a deeper insight in pUL5 role during infection, an *UL5* deletion mutant was generated by BAC mutagenesis and its ability to replicate in human fibroblasts was assessed. The knock-out mutant virus is still able to infect and replicate in fibroblasts but with a strongly reduced replication rate, suggesting an important role of pUL5 during HCMV infection. The recombinant pUL5-tagged virus was also implemented for a pull-down experiment to isolate potential interacting partners using anti-FLAG coupled magnetic beads. Mass spectrometry analysis allowed the identification of the cellular protein IQGAP1 as a pUL5 interactor. IQGAP1 is an ubiquitously expressed cellular protein involved in many cellular mechanisms,

such as cytoskeletal remodeling, cell adhesion, cell cycle regulation and Ca^{2+}/Calmodulin signaling, and it is a known target of intracellular pathogens. This interaction was also confirmed by reciprocal immunoprecipitations in transfection, giving new hints on the possible role of pUL5 during HCMV infection.

CHAPTER I

INTRODUCTION

HUMAN CYTOMEGALOVIRUS GENERAL FEATURES, PATHOGENESIS AND CLASSIFICATION

Human cytomegalovirus (HCMV) is a ubiquitous virus infecting the majority of adults worldwide, ranging around 55-60% of the world population, although reaching peaks of more than 90% seroprevalence in given populations grouped according to age-gender-socioeconomic related factors [1]. Horizontal transmission depends on direct contact of infected bodily fluid at mucosal surfaces. In healthy individuals, a strong immune response to HCMV is able to limit and contain the spread of the disease, therefore bland clinical symptoms and spontaneous acute infection resolution are associated to infection of immunocompetent individuals [1]. Nevertheless, the virus is able to establish a lifelong lasting latency with recurrent and spontaneous reactivation and complete clearance of the virus by the organism is never achieved. Both innate and adaptive immune responses operate to control viral replication and spread. The acute disease associated with HCMV infection occurs only in a small proportion of infected individuals, in particular its onset is observed in patients where the normal immune response is compromised. For example, HCMV infection can lead to severe complications and death in the case of transplant patients when the immune system is shut down to avoid rejection of the organ [2]. Another example of opportunistic superinfection of HCMV is very common in patients with Acquired Immuno-Deficiency Syndrome (AIDS) following HIV infection; the immuno-deficient environment in these patients allows HCMV infection and/or reactivation [3]. Moreover, due to the HCMV ability to cross the trans-placental barrier, congenitally acquired infection can pose a severe threat to the fetus that lacks a fully functional immune system. In all these particular cases, antiviral therapies are

necessary to counteract the severe disease associated to the HCMV primary infection or reactivation. Unfortunately, almost all antiviral agents actually on the market show a significant number of potential side effects including organ and marrow toxicity and the onset of increasing drug resistant viral strains [1]. To reduce the morbidity and mortality associated to HCMV infection in immunocompromised patients, a commonly implemented prophylactic tool is the passive immunization using hyper-immune globulin preparations with high levels of antibodies against HCMV (CMVIG). For example, combination of CMVIG and the antiviral drug ganciclovir proved to be effective in preventing HCMV infection in patients undergoing solid organ transplant [4]. Moreover, a study suggests their effectiveness in the treatment and prevention of severe diseases associated to congenital HCMV infection [5]. Nevertheless, the effectiveness and efficacy of CMVIG seem to be restricted to specific clinical cases while only modest beneficial effects were observed upon CMVIG treatment in other settings such as blood transfusion in premature newborns and hematopoietic cell transplantation (HCT) [6]. Some CMVIGs exhibit low potency in vitro [7] and must be administrated at high doses to exert a beneficial effect. To overcome this problem, a series of neutralizing monoclonal antibodies against immune-dominant HCMV envelope glycoproteins were developed. For example, a highly in vitro neutralizing antibody against glycoprotein H was isolated from the spleen of a CMV seropositive patient. The monoclonal antibody, named MSL-109 or sevirumab, was tested in phase II clinical trials as treatment for CMV-induced retinitis in AIDS patients [33]. Unfortunately the tests were stopped due to lack of efficacy [8].

HCMV belongs to the *Herpesviridae* family, a classification historically based on the architecture of the virion that includes the herpesviruses of mammals, reptiles and birds. The *Herpesviridae* family is one of the three families, together with the *Alloherpesviridae* and the *Malacoherpesviridae,* encompassed by the *Herpesvirales* order, established later as extensive nucleotide sequence data became available [9]. The *Herpesviridae* family members share several biological properties, the most evident is the ability to establish latent infection in some cell types that usually vary among the different members. Moreover, in addition to

similar structural features, they share a large array of enzymes involved in nucleic acid metabolism, DNA synthesis and protein processing. Finally, in all *Herpesviridae* the DNA replication and capsid assembly occur in the nucleus, the final processing of the virion takes place in the cytoplasm and the release of the progeny results in the destruction of the host cell [1]. Among the *Herpesviridae*, the most common ones having human cells as their primary reservoir are: herpes simplex virus (HSV-1 and HSV-2), HCMV, varicella-zoster virus (VZV), Epstein-Barr virus (EBV), and *Human herpesviruses 6*, *7*, and *8* (HHV-6, HHV-7, and HHV-8). The *Herpesviridae* family is itself composed of three subfamilies: the *Alphaherpesvirinae* (HSV-1, VZV), with a relatively short reproductive cycle and the ability the establish latent infection primarily in sensory ganglia, the *Betaherpesvirinae* (HCMV, MCMV), with a long reproductive cycle resulting in enlarged infected cells, and the *Gammaherpesvirinae* (EBV), usually specific for B and T lymphocytes and latent in the lymphoid tissue [10]. All betaherpesviruses exibit species specificity and only particular types of differentiated cells are susceptible within a species. All CMV are conveniently propagated in fibroblasts although natural targets include other cell types, such as epithelial, endothelial, macrophage and dendritic cells [11]. Unexpectedly, tropism is not determined by the distribution of cellular receptors on the host cells, in fact, non-permissive cells generally allow viral attachment and penetration, but viral replication is blocked at an early post-entry step [1].

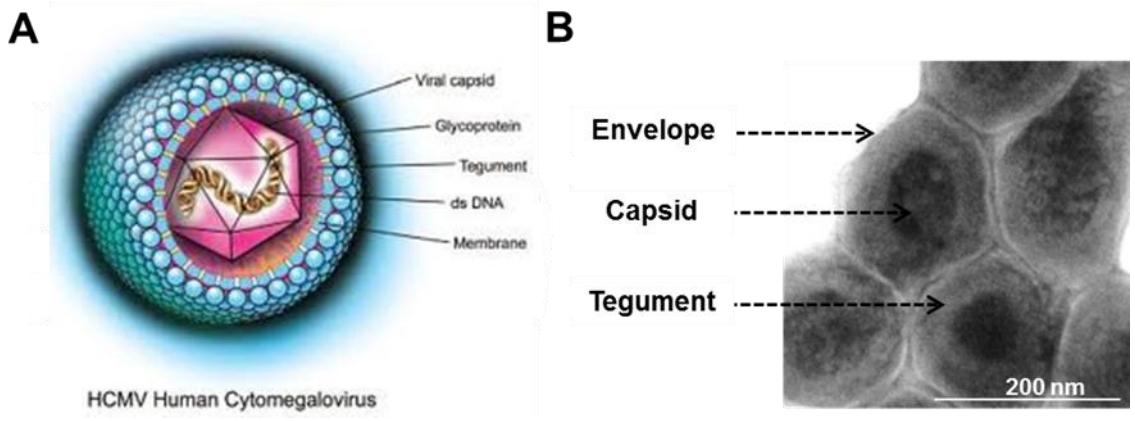

Figure 1. (A) Schematic representation of HCMV virion (from http://drmhanna.com/cytomegalovirus-cmv/). (B) Negative stained purified virion preparation analyzed through electron microscopy.

The HCMV virion is composed of an envelope, a tegument, and a capsid containing the double-stranded DNA genome. The virus envelope is a lipid bilayer derived from the host cell endoplasmic reticulum (ER)-Golgi intermediate compartment (ERGIC) and contains the viral glycoproteins necessary for the cellular tropism [12]. Most functionally critical HCMV envelope glycoproteins are common to all herpesviruses and form three important complexes, termed *gcI* (glycoprotein B - gB - homodimer), *gcII* (gH:gL complex) and *gcIII* (gM:gN complex) [12-15]. Furthermore, two additional complexes of gH:gL have been observed, one including gO and probably involved in gH:gL-mediated fusion [16] and the other is a pentameric complex with *UL128*, *UL130* and *UL131* gene products and seems to be responsible for the specific tropism of some HCMV clinical isolates for epithelial, endothelial and dendritic cells [17, 18]. In comparison to the other herpesviruses, the envelope is highly irregular resulting in a viral shape whose diameter reaches up to 300 nm with an average size of 230 nm in diameter [19].

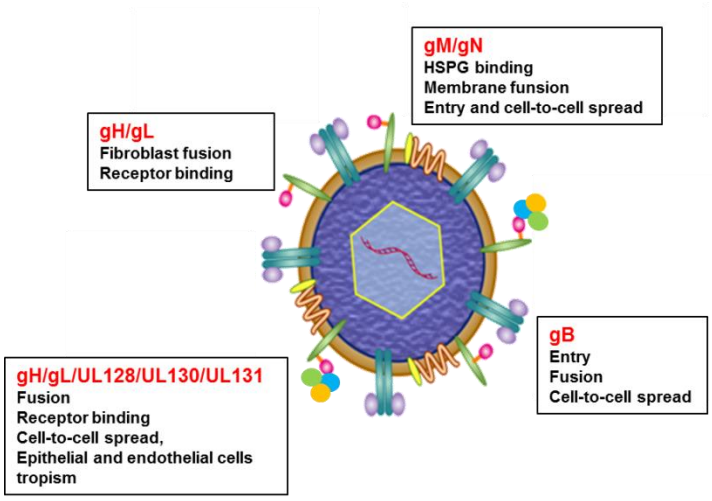

Figure 2: HCMV envelope glycoprotein complexes. The proposed role of each complex is indicated in the respective box. Modified from [20].

The layer beneath the envelope is an amorphous matrix referred to as tegument. This compartment carries the majority of the HCMV expressed proteins plus a huge number of host species including proteins and nucleic acids. Tegument proteins are usually phosphorylated [13, 21] and cover several crucial viral functions including disassembly of the viral particle during entry, modulation of the host cell response to infection, alteration of the host cell cycle, and assembly of new particles, all aiming to the creation of an optimal intracellular environment for viral replication; among these are for example kinases, trans-activating factors and the most abundant and immunogenic protein pp65 [22].

The HCMV capsid is embedded in the tegument and contains a tightly packed linear double-stranded DNA genome. The complex of capsid and DNA is commonly named nucleocapsid. The icosahedral nucleocapsid is composed of five herpesvirus core proteins: the major capsid protein (MCP), the minor capsid protein and minor capsid protein binding protein (TRI1 and TRI2 respectively), the smallest capsid protein (SCP), and the portal protein (PORT) responsible for the encapsidation of the viral DNA [13]. Within infected cells, three mature capsid forms are observed, A, B, and C capsids. Only the latter type contains viral

DNA, while A and B capsids are usually found in the nucleus and cytoplasm of infected cells and are commonly referred to as non-infectious enveloped particles (NIEPs) [12, 13].

In line with the huge virus dimension, HCMV has the largest genome among its family with approximately 230 kb, 50% larger than HSV, despite the capsids of both viruses being roughly of the same size (110-125 nm) [23, 24]. The genome organization reflects the common structure of all herpesviruses, composed of unique long (U_L) and unique short (U_S) genetic regions flanked by two sets of inverted repeats (R_L, repeated long and R_S, repeated short). Recombination phenomenon can occur among identical terminal and internal repeats, leading to genome isomerization. Thus, genetic material isolated from a viral population consists of equal amounts of four different genomic isomers pooled together [1]. While the general genomic arrangement is conserved among HCMV strains, a major difference in the open reading frame (ORF) organization and composition can be observed between "laboratory-adapted strains" and "clinical isolates" [25]. While laboratory strains indicate all the strains extensively passaged and adapted to growth in human fibroblasts, permitting easy manipulation due to the fast replication and high yields of produced virus, clinical isolates underwent through none or limited passages in cell culture before being cloned as bacterial artificial chromosomes (BACs) and/or sequenced. As a consequence of the fibroblast adaptation, severe genomic rearrangements occurred in laboratory adapted strains. Both Towne and AD169 laboratory strains acquired a large deletion of the U_L segment, concomitantly being replaced by a duplication of the R_L region. Due to these rearrangements, a difference in the coding potential can be observed between laboratory strains and clinical isolates. In particular, while the AD169 is predicted to encode for 208 ORFs (including the repeated segments), the coding potential of a clinical isolate is estimated around 252 ORFs [25, 26]. Recently, through an integrated approach that combines ribosome footprinting assays and high-resolution mass spectrometry, Stern-Ginossar and co-workers reported no less than 751 translated ORFs and 53 novel proteins originating from ORFs overlapping the known ORFs [27]. Despite this huge coding potential, only a small subset of these proteins constitutes the

mature virion: it has been estimated that around 50 [1, 28] proteins are incorporated in the virion.

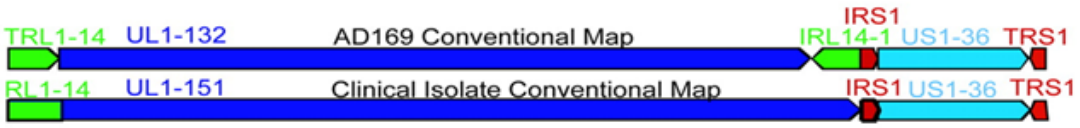

Figure 3. HCMV genome organization. ORFs map of conventional laboratory strain AD169 and clinical isolate. The AD169 genome (upper) carries TRL1-14 (green arrow), UL1-132 (dark blue arrow), IRL14-1 (green arrow), IRS1 (red arrow), US1-36 (light blue arrow), and TRS1 (red arrow). In the clinical isolate RL1-14, (green arrow), UL1-151 (dark blue arrow), IRS1 (red arrow), US1-36 (light blue arrow), and TRS1 (red arrow) are present. Modified from [29].

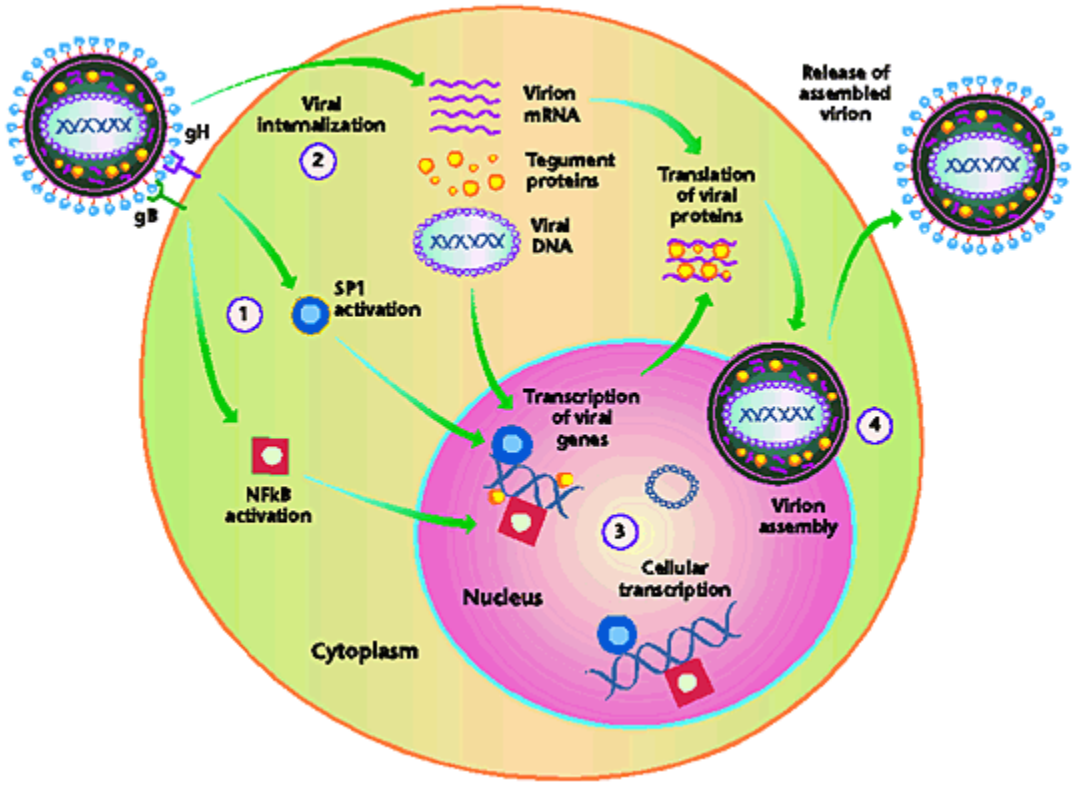

Figure 4. HCMV lytic life cycle. (1) Binding of HCMV glycoproteins gB and gH to cellular receptors activates cellular transcription factors, such as NF-kappaB and Sp1. (2) The virus then enters the cell, releasing viral DNA, virion proteins and virion mRNA transcripts into the cytoplasm, where virion mRNAs are translated. Viral DNA and certain viral proteins are transported to the nucleus. (3) In the nucleus, viral and cellular genes are expressed, with help from the activated transcription factors, and viral DNA is replicated. (4) Viral DNA, viral and cellular proteins, and virion transcripts are packaged into the virion through an as yet undetermined mechanism. During egress of the virion from the cell, the virion envelope is constructed and an infectious viral particle is released. From [30].

HCMV replicates in many cell types, but most of the available information originates from studies conducted in primary or secondary fibroblasts, a stromal cell type that can be derived from any host tissue. In general, three main classes of gene expression determine the productive infection of HCMV in permissive cells: immediate early (IE), delayed-early (DE) and late (L) genes. Despite the slow replication cycle, requiring 48 to 72 hours for the progeny to egress, the expression of IE gene products begins immediately after viral entry [1]. This initial step consists in five consequential events: the binding of viral envelope glycoproteins to

specific cell surface receptors, the release of nucleocapsid into the cytoplasm mediated by the fusion of the viral envelope with the cellular membrane, the nucleocapsid association with cytoskeletal elements to promote translocation towards the nucleus, the nucleocapsid interaction with the nuclear pores and the release of the viral genome into the nucleus.

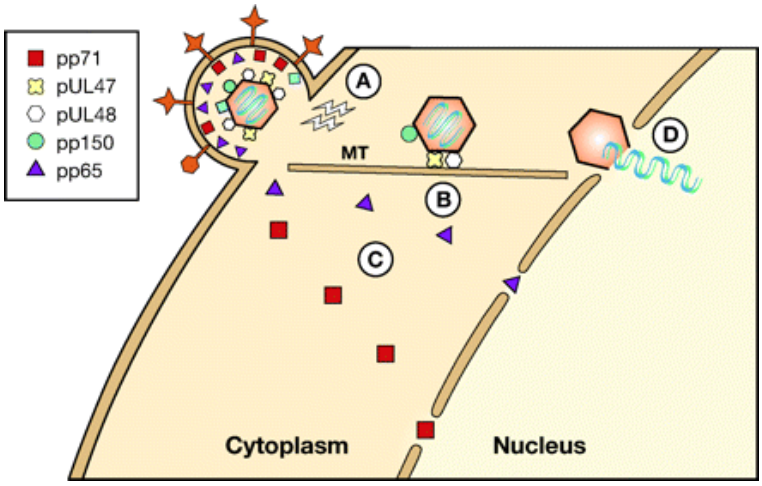

Figure 5: Delivery of HCMV capsid to the nucleus. **(**A) HCMV capsid and associated tegument proteins are transported along the microtubules (B) toward the nucleus. A set of tegument proteins dissociate from the capsid and migrates independently to the nucleus (C). Viral DNA is released into the nucleus through the nuclear pores (D). Modified from [31].

A relatively conserved feature among the herpesviral family is the engagement of heparin sulfate proteoglycans (HSPG) that is believed to initiate a cascade of events involving other entry mediators and cellular receptors [14]. The principle actor of this initial binding in HCMV infection is the gB homodimer. Apparently, the eterodimeric complex composed of the most abundant envelope glycoprotein, gM, and gN is also involved in HSPG tethering and stabilization of virus-cell contact. The gH-gL complex plays an important role in the fusion of the viral particle with the cell membrane and variable third components appear to be responsible for different cell tropisms: gO for fibroblasts and pUL128, pUL130 and pUL131 for epithelial and endothelial cells and monocyte-macrophages tropism [32]. Various works describing the HCMV entry step suggest that the virus uses distinct cellular receptors, and consequently different entry pathways, depending on the target cell. Indeed, while direct

fusion at neutral pH has been observed *in vitro* in fibroblast cells [14], low pH dependent receptor-mediated endocytosis is required for viral entry in epithelial and endothelial cells [33]. Once the particle is released into the cytoplasm, regulatory tegument proteins dissociate from the capsid and either remain in the cytoplasm or migrate independently to the nucleus, where they modulate cellular and viral genes expression. A subset of tegument proteins remains tightly associated with the nucleocapsid and contributes to the delivery of the DNA to the nucleus by association with elements of the microtubular network (MT). HCMV nucleocapsids move from cellular periphery along MT branches up to the perinuclear MT organizing center (MTOC) to reach the nuclear pore complex and to inject the viral genome into the nucleus [34]. Almost immediately after the viral DNA has been delivered to the nucleus, IE genes expression ensues resulting in dramatic effects on the host cells. On the one hand, IE gene products act as global regulators of viral genes expression but they also promote several functions, such as cell death suppression or immune-evasion, important for viral pathogenesis. Delayed-early genes expression starts by 6 h post-infection and continues during viral DNA synthesis. Indeed, the majority of DE gene products are involved in viral DNA replication and in its fine regulation [35], while others are implicated in capsid maturation and modulation of the host cell response to infection. During this phase of infection, a gradual increase in HCMV DNA synthesis and a strong stimulation of the host metabolism are observed, while, at least in fibroblasts, the host cell cycle is blocked and no cellular DNA synthesis occurs [1]. Furthermore, many DE genes, together with L genes, encode for structural proteins or are involved in the final steps of the viral replication cycle, such as capsid maturation, DNA encapsidation and virion maturation and egress from the host cell. HCMV infection induces a global reprogramming of the cellular activity and causes dramatic changes in infected cell morphology. The most evident is the formation of two compartments: a nuclear replication inclusion where the viral DNA synthesis takes place and a large cytoplasmic juxtanuclear region defined virion assembly compartment (AC), responsible for the virus final envelopment and egress. Moreover, HCMV has developed a huge arsenal of

genetic functions committed to modulate both innate and adaptive immune responses. In particular, the virus is able to subvert the immune system mimicking the same strategies and mechanisms used by the host cells to clear the infection. Inhibition of complement cascade and natural killer (NK) cells activation, attenuation of interferon (INF) response and disruption of antigen presentation are only few examples of the functions hijacked by virus encoded chemokines, cytokine and cellular receptors homologues [36].

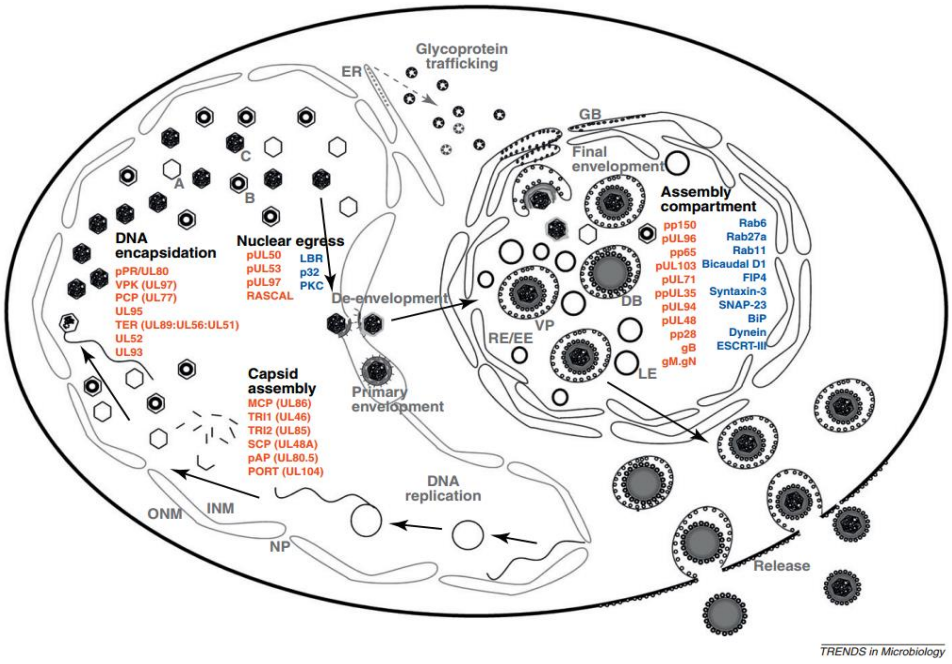

Figure 6. HCMV virion particles formation, maturation and budding processes. Major cellular and viral proteins involved in these processes are reported (red and blue, respectively). List of abbreviations: DB, dense body; VP, virus particle; EE/RE, early endosome/recycling endosome; LE, late endosome; GB, Golgi body; ER, endoplasmic reticulum; NP, nuclear pore; INM, inner nuclear membrane; ONM, outer nuclear membrane; A, B and C, types of nuclear capsids. From [29].

The final steps are common among the *Herpesviridae* family and involve a two-stage envelopment-deenvelopment-envelopment process that drives the nucleocapsid outside the nucleus to the AC to acquire the full spectrum of tegument proteins and envelope glycoproteins, and finally the mature virion to be released by exocytosis at the plasma membrane [12].

REFERENCES

1. Mocarski, E.S., T. Shenk, and R.F. Pass, Cytomegaloviruses. In: Knipe DM, Howley PM, Fields Virology. Wolters Kluwer, Lippincott Williams & Wilkins. 2013.

2. van der Bij, W. and R. Speich, *Management of Cytomegalovirus Infection and Disease after Solid-Organ Transplantation.* Clinical Infectious Diseases, 2001. **33**(Supplement 1): p. S32-S37.

3. Cheung, T.W. and S.A. Teich, *Cytomegalovirus infection in patients with HIV infection.* Mt Sinai J Med, 1999. **66**(2): p. 113-24.

4. Zamora, M.R., et al., Following universal prophylaxis with intravenous ganciclovir and cytomegalovirus immune globulin, valganciclovir is safe and effective for prevention of CMV infection following lung transplantation. Am J Transplant, 2004. **4**(10): p. 1635-42.

5. Nigro, G., et al., *Passive Immunization during Pregnancy for Congenital Cytomegalovirus Infection.* New England Journal of Medicine, 2005. **353**(13): p. 1350-1362.

6. !!! INVALID CITATION !!!

7. Roy, D.M. and J.E. Grundy, Evaluation of neutralizing antibody titers against human cytomegalovirus in intravenous gamma globulin preparations. Transplantation, 1992. **54**(6): p. 1109-10.

8. Jabs, D.A., et al., HIV and cytomegalovirus viral load and clinical outcomes in AIDS and cytomegalovirus retinitis patients: Monoclonal Antibody Cytomegalovirus Retinitis Trial. Aids, 2002. **16**(6): p. 877-87.

9. McGeoch, D.J., F.J. Rixon, and A.J. Davison, *Topics in herpesvirus genomics and evolution.* Virus Res, 2006. **117**(1): p. 90-104.

10. Roizman, B., et al., Herpesviridae. Definition, provisional nomenclature, and taxonomy. The Herpesvirus Study Group, the International Committee on Taxonomy of Viruses. Intervirology, 1981. **16**(4): p. 201-17.

11. Gerna, G., F. Baldanti, and M.G. Revello, *Pathogenesis of human cytomegalovirus infection and cellular targets.* Hum Immunol, 2004. **65**(5): p. 381-6.

12. Britt, B., *Maturation and egress*, in *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, A. Arvin, et al., Editors. 2007, Cambridge University Press

Copyright (c) Cambridge University Press 2007: Cambridge.

13. Britt, W.J. and S. Boppana, *Human cytomegalovirus virion proteins.* Hum Immunol, 2004. **65**(5): p. 395-402.

14. Compton, T. and A. Feire, Early events in human cytomegalovirus infection, in Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis, A. Arvin, et al., Editors. 2007, Cambridge University Press

Copyright (c) Cambridge University Press 2007: Cambridge.

15. Dunn, W., et al., *Functional profiling of a human cytomegalovirus genome.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14223-8.

16. Paterson, D.A., et al., A role for human cytomegalovirus glycoprotein O (gO) in cell fusion and a new hypervariable locus. Virology, 2002. **293**(2): p. 281-94.

17. Wang, D. and T. Shenk, *Human cytomegalovirus virion protein complex required for epithelial and endothelial cell tropism.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(50): p. 18153-18158.

18. Hahn, G., et al., Human Cytomegalovirus UL131-128 Genes Are Indispensable for Virus Growth in Endothelial Cells and Virus Transfer to Leukocytes. Journal of Virology, 2004. **78**(18): p. 10023-10033.

19. Liu F and Z. ZH, Comparative virion structures of human herpesviruses. In: Arvin A, Campadelli-Fiume G, Mocarski E, et al., editors. Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. Cambridge University Press; Chapter 3. Available from: http://www.ncbi.nlm.nih.gov/books/NBK47399/, 2007.

20. Compton, T., Receptors and immune sensors: the complex entry path of human cytomegalovirus. Trends Cell Biol, 2004. **14**(1): p. 5-8.

21. Munger, J., D. Yu, and T. Shenk, UL26-deficient human cytomegalovirus produces virions with hypophosphorylated pp28 tegument protein that is unstable within newly infected cells. J Virol, 2006. **80**(7): p. 3541-8.

22. Tomtishen Iii, J.P., Human cytomegalovirus tegument proteins (pp65, pp71, pp150, pp28). Virology Journal, 2012. **9**: p. 22-22.

23. Davison, A.J., et al., The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. Journal of General Virology, 2003. **84**(1): p. 17-28.

24. Dolan, A., et al., *Genetic content of wild-type human cytomegalovirus.* Journal of General Virology, 2004. **85**(5): p. 1301-1312.

25. Murphy, E., et al., *Coding potential of laboratory and clinical strains of human cytomegalovirus.* Proceedings of the National Academy of Sciences, 2003. **100**(25): p. 14976-14981.

26. Murphy, E., et al., *Reevaluation of human cytomegalovirus coding potential.* Proceedings of the National Academy of Sciences, 2003. **100**(23): p. 13585-13590.

27. Stern-Ginossar, N., et al., *Decoding Human Cytomegalovirus.* Science, 2012. **338**(6110): p. 1088-1093.

28. Varnum, S.M., et al., Identification of Proteins in Human Cytomegalovirus (HCMV) Particles: the HCMV Proteome. Journal of Virology, 2004. **78**(20): p. 10960-10966.

29. Britt, W., Manifestations of human cytomegalovirus infection: proposed mechanisms of acute and chronic disease. Current topics in microbiology and immunology, 2008. **325**: p. 417-470.

30. Huang, E.-S. and R.A. Johnson, *Human cytomegalovirus - no longer just a DNA virus.* Nat Med, 2000. **6**(8): p. 863-864.

31. Kalejta, R.F., *Tegument proteins of human cytomegalovirus.* Microbiol Mol Biol Rev, 2008. **72**(2): p. 249-65, table of contents.

32. Schauflinger, M., et al., Analysis of human cytomegalovirus secondary envelopment by advanced electron microscopy. Cell Microbiol, 2013. **15**(2): p. 305-14.

33. 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.

34. Döhner, K., *The role of the cytoskeleton during viral infection.* Current Topics in Microbiology and Immunology, 2005. **285**: p. 67-108.

35. Shikhagaie, M., et al., The Human Cytomegalovirus-Specific UL1 Gene Encodes a Late-Phase Glycoprotein Incorporated in the Virion Envelope. Journal of Virology, 2012. **86**(8): p. 4091-4101.

36. Miller-Kittrell, M. and T.E. Sparer, *Feeling manipulated: cytomegalovirus immune manipulation.* Virol J, 2009. **6**: p. 4.

CHAPTER II

CHARACTERIZATION OF pUL5, AN HCMV PROTEIN INTERACTING WITH THE CELLULAR PROTEIN IQGAP1

ABSTRACT

HCMV is a β-herpesvirus infecting the majority of population, usually establishing a lifelong latent infection. Immunosuppressive conditions, such as that of transplant, AIDS patients or newborns, allow reactivation of the virus, associated to a variety of diseases. Among *herpesviridae*, HCMV owns the largest genome and displays a huge coding potential. In this work, we address the study of the gene product of *UL5* reading frame of the clinical isolate TR strain, present as a single copy in both clinical isolates and laboratory adapted HCMV genomes. The putative protein product is a 166-amino-acid membrane protein with a theoretical mass of 19 kDa. Northern blot analyses showed that the gene was transcribed with early kinetics together with *UL4* reading frame. Recombinant virus expressing pUL5 with a tag at the C-terminus of the protein allowed the identification of two pUL5 non-glycosylated species of approximately 24 and 15 kDa, expressed with late and early kinetics, respectively. Experiments in transfection suggest that the lower molecular weight species could be the result of the translation from a second ATG in *UL5* reading frame. Confocal experiments carried out both with the recombinant and with the wild-type virus localize pUL5 in the assembly compartment, but western blot on purified virions failed to identify the protein, suggesting that pUL5 is only expressed in the host cell during infection. Furthermore, *UL5* deletion resulted in an impaired replication of the virus in human fibroblasts with respect to the wild-type virus. Finally, pull-down experiments coupled with mass spectrometry analysis identified IQGAP1 as a pUL5 interactor, and this was also confirmed by reciprocal

immunoprecipitations in transfection, giving new hints on pUL5 possible roles during HCMV infection.

INTRODUCTION

HCMV is a ubiquitous, highly specific herpes virus, infecting as many as 75% of adults. It establishes latent infection adapting numerous immune evasion strategies. Despite being normally relatively benign, HCMV infection can induce severe disease in immunocompromised patients, and in fetuses in the case of primary maternal infection [1]. HCMV infects a wide spectrum of cell types using different sets of surface molecules. This heterogeneity suggests a very complex mechanism of recognition between the viral particle and the host cell. To obtain deeper insights of the mechanisms involved in HCMV infection, tropism, immune evasion and latency, it is yet required to study HCMV proteome.

HCMV displays a coding capacity that far exceeds that of most other *Herpesviridae*, having the largest genome among all known human viruses. Its 240-kb double-stranded linear DNA genome is composed by a long unique (U_L) and a short unique (U_S) sequences, each flanked by inverted repeats, named terminal or internal based on their position (TR_L/IR_L and TR_S/IR_S), giving the overall genome configuration TR_L-U_L-IR_L-IR_S-US-TR_S [2]. Within the *Herpesviridae* family, sequence variability is very high but the gene content is quite conserved, with a set of core genes involved in metabolic and structural functions shared by all members. On the other hand, each member of this heterogeneous virus family carries some private genes. In most cases, these genes are dispensable for viral replication in cell culture and are usually involved in viral tropism or in the evasion from the immune system. Among them, the RL11 gene family includes 11 genes present only in human and chimpanzee cytomegaloviruses, and 3 genes only present in HCMV [3]. A common characteristics of this family is the RL11 domain (RL11D), an immunoglobulin-like domain shared by the majority of the members of the RL11 family [4]. Belonging to this family are: two receptors for the Fc domain of immunoglobulin G, RL11 and RL13 [5, 6], three are envelope-associated glycoproteins (*UL4*-encoded gp48, UL1 and RL13) [7-9]. The product of *UL11* gene is a hypervariable protein expressed on the surface of infected cells [10, 11]. *UL7* encodes for a SLAM-family receptor

CD229 homologue that impairs cytokine production and promotes angiogenesis [12, 13]. In general, RL11 family members share some common features, in particular that of membrane-associated glycoproteins and the RL11D with some exceptions: UL4, RL5 and RL6 for example lack the transmembrane region [3, 8].

In the present study, we have approached the structural and functional characterization of *UL5*, member of the RL11 family that lacks the RL11D and appears to be N-terminally truncated with respect to the other RL11 family members. To this goal, we produced a recombinant HCMV virus adding a combination of tags (two StrepII and two FLAG tags) at the C-terminus of UL5 protein product by BAC mutagenesis of the TR strain genome of HCMV in *E. coli*, and a deletion mutant lacking the entire *UL5* open reading frame (ORF). Our data indicate that pUL5 is an early expressed protein that is not incorporated into the viral particle but localizes in the assembly compartment of the infected cell. Furthermore, we demonstrated that *UL5* deletion results in impaired replication in human fibroblasts and that during infection it interacts with the cellular scaffold protein IQGAP1.

RESULTS AND DISCUSSION

UL5 OPEN READING FRAME AND ITS PREDICTED PROTEIN PRODUCT

The *UL5* ORF is found as a single copy in the HCMV genome, within the unique long region, approximately 2 kb from the junction with the terminal repeat long region. According to sequence similarity, *UL5* has been classified as a member of the RL11 multigene family [4]. This family consists of 14 genes (*RL5A, RL6, RL11, RL12, RL13, UL1, UL4, UL5, UL6, UL7, UL8, UL9, UL10, UL11*) distributed contiguously in the long repeats and in the unique long region, aligned from 5' to 3' orientation with this order. There are two additional ORFs (*UL2* and *UL3*) identified on the opposing strand between *UL1* and *UL4* as unrelated genes [3]. The RL11 family is defined by the RL11 domain, a region of variable length with three conserved residues (a tryptophan and two cysteines) that resembles a domain shared by some members of the immunoglobulin superfamily [4]. Despite being an RL11 family member, *UL5* lacks an RL11 domain and is N-terminally truncated but does have a transmembrane domain. All RL11 genes are absent in cytomegaloviruses of other animal species, with the exception of chimpanzee cytomegalovirus (CCMV) that contains counterparts of all but *UL1* [3]. *UL5*, therefore, is only present in human and chimpanzee cytomegaloviruses while is absent in cytomegaloviruses infecting other species [2].

In order to investigate *UL5* ORF and the similarity of amino-acid sequence among different HCMV strains, a multiple alignment was performed using CLUSTALW server [14]. In the alignment both laboratory adapted strains and clinical isolates are compared, showing that the overall sequence is conserved and major differences belong to the N-terminal part of the predicted protein (Fig. 1A).

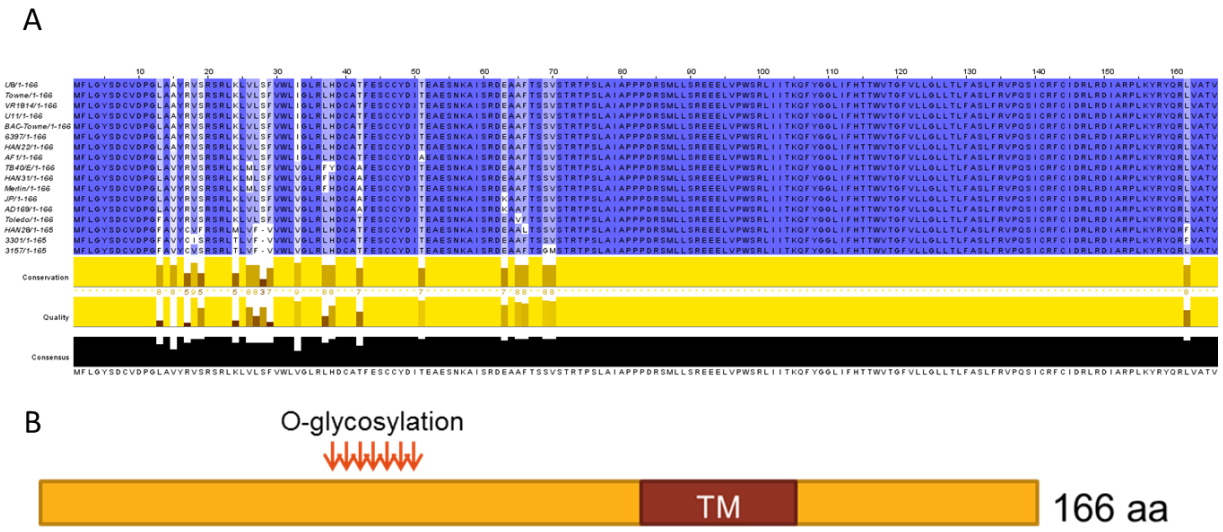

Figure 1. (A) Alignment of *UL5* amino acid sequences of clinical isolates and laboratory adapted strains. Numbers above the sequences represent amino acid residues. (B) Structural prediction of HCMV UL5 protein. Transmembrane region and predicted O-glycosylation sites are indicated.

UL5 ORF is predicted to encode a 166 amino-acid protein. Analysis with PSORT server [15] identified a transmembrane region between amino-acids 118 and 134, suggesting that pUL5 is a type Ib membrane protein with the C-terminus exposed in the cytoplasmic region and the N-terminus exposed towards the lumen or extracellular space. PSORT analysis also excluded the presence of a signal peptide. This prediction was further corroborated by the SIGNALP server prediction [16]. Since many RL11 family members are predicted or have been identified as membrane glycoproteins [6, 9-11], the glycosylation prediction was also investigated using NetNGlyc [17] and NetOGlyc [18] servers. pUL5 is not predicted to have N-glycosylation sites while 7 potential sites of O-glycosylation are present (Fig. 1B).

According to literature, *UL5* coding sequence is included in three bigger transcripts containing also *UL4* ORF and transcribed from three promoters with different kinetics [8, 19]. *UL4* protein product, also called gp48, is a glycoprotein expressed in the early phase of infection that seems to have a very complex regulation, at the transcription level as well as the translation, due to an upstream ORF present in the same transcripts [20, 21]. To date, no information is available on the protein product of *UL5* ORF.

In order to investigate the transcription profile of the *UL5* ORF, total RNA was isolated from human foreskin fibroblast (HFF) cells at different time points after infection with HCMV TR and subjected to Northern blot analysis using a *UL5* mRNA-specific biotinylated DNA probe. The probe detected a single transcript of 1.5 kb already visible at 6 hours post infection (p.i.) (Fig. 2). This is consistent with the information available in literature, confirming that *UL5* mRNA is part of a bicistronic transcript with *UL4* and excluding the presence of another transcript containing *UL5* alone. In fact, also the promoter prediction analysis failed to identify putative promoters between *UL4* and *UL5* coding sequences (data not shown).

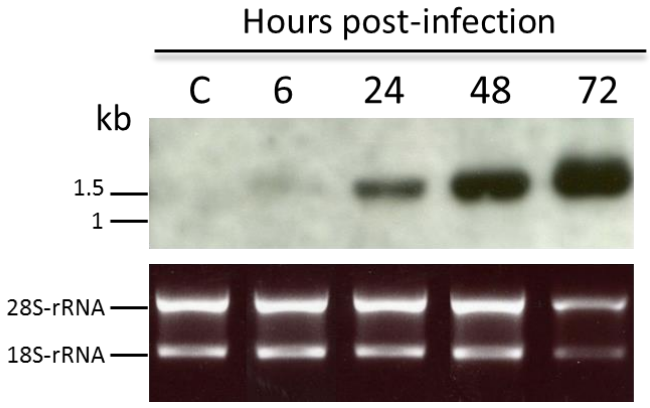

Figure 2. *UL5* gene transcription. Northern blot of whole-cell RNA isolated from uninfected (C) or TR infected fibroblasts at the indicated times post-infection with a *UL5*-specific probe.

CONSTRUCTION OF UL5-TAGGED RECOMBINANT VIRUS

In order to investigate UL5 protein product in the contest of infection, we applied a two-step Red-Gam mediated mutagenesis on a BAC containing the entire genome of HCMV TR strain [22] to obtain a recombinant virus carrying a tag at the C-terminus of UL5 protein. Briefly, this strategy consists of two recombination steps: the first allows the integration of the tag together with the expression cassette for kanamycin resistance, while in the second step, digestion with I-SceI enzyme permits the alignment of homologous sequences and the excision of the kanamycin resistance cassette, leaving the tag in-frame with the C-terminus of pUL5. This is achieved with the implementation of the *E. coli* strain GS1783 containing the TR BAC, the lambda Red system expressed with a heat-inducible promoter and the I-SceI enzyme expression cassette under the control of a promoter inducible with arabinose.

Initially, this strategy was implemented with a cassette amplified from pEP-KanS vector that allows the adding of a FLAG tag at the C-terminus of the protein of interest. After several attempts, we concluded that this was not the optimal tag for UL5 protein detection and, therefore, we constructed a new cassette for recombination that could permit the modification of the protein of interest with multiple combinations of tags in order to identify the best tag for UL5 protein. This cassette was designed in a palindrome manner (Fig. 3), so that the amplification of the fragment indicated by the upper arrow would result in the adding of V5 and streptavidin-binding peptide (SBP) tags, while the amplification of the opposing strand as indicated by the lower arrow would add two StrepII and two FLAG tags at the C-terminus of the protein of interest. Furthermore, in addition to the possibility of testing different tags for UL5 protein detection, this strategy allows also the implementation of the tandem-affinity purification (TAP), i.e. the sequential purification of a protein by affinity binding to different tags, in order to isolate UL5 protein together with its interacting partners. In fact, it has been shown that these combinations of tags have given the best results when used for mammalian proteins immunoprecipitations [23].

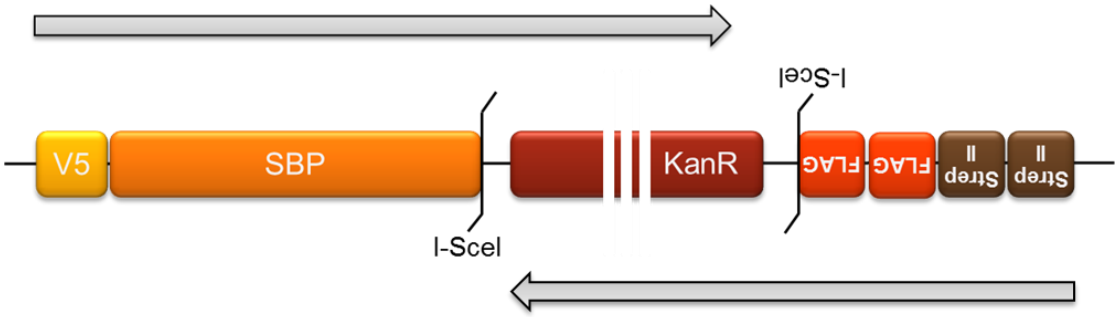

Figure 3. Recombination cassette of pEP-TAP. The cassette is designed for the differential tagging of a selected protein, giving the chance to introduce V5 and SBP tags if amplified as indicated by the upper arrow or two StrepII and two FLAG tags if amplified as shown by the lower arrow.

UL5 PROTEIN EXPRESSION

The expression of UL5 protein product during HCMV infection was investigated using the recombinant virus with a tagged-UL5 protein. To this purpose recombinant infectious viruses were used to infect HFF cells at multiplicity of infection (MOI 5). Lysates were prepared at different times post-infection (p.i.) and subjected to immunoblot using anti-tag monoclonal antibodies to assess the presence of UL5. The best signal was obtained using the anti-FLAG monoclonal antibody on lysates of TR-UL5-2StrepII-2FLAG infected HFF cells (Fig. 4). HCMV lytic gene expression is conventionally divided into three major kinetic classes of viral genes, immediate-early, early, and late. Monoclonal antibodies against HCMV immediate-early antigen 1 (IE1) and gB (late gene) were used as references for the expression kinetics and GAPDH as a loading control. Two species of approximately 24 and 15 kDa were detected starting 48 h p.i with different expression patterns. The lower molecular weight species is already fully expressed at 48 h p.i. while the higher is less expressed. Both species are absent at 24 h p.i. while IE1 is already visible, demonstrating that pUL5 is not expressed with an immediate-early kinetics. Apparently, the higher molecular weight form of pUL5 shows an expression profile similar to that of gB, consistent with a late expression profile, while the lower molecular weight form appear to be expressed with an early kinetics.

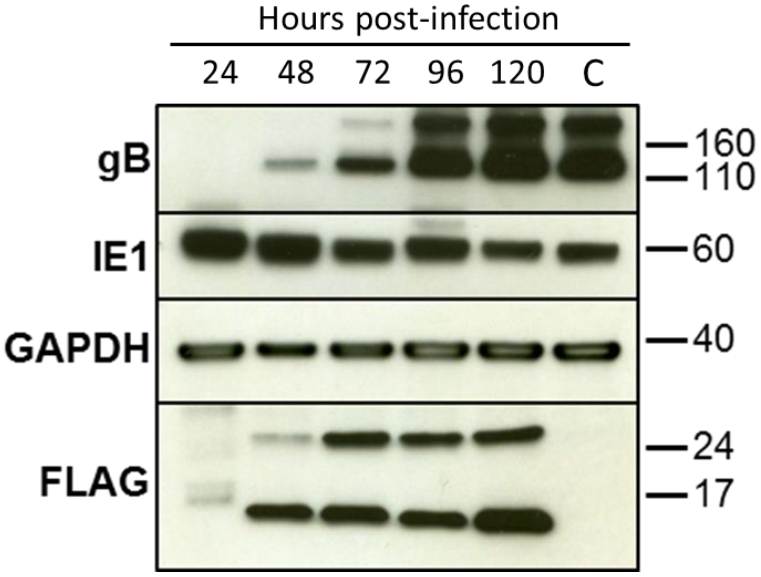

Figure 4. Expression of UL5 protein in HCMV-infected fibroblasts. Western blot on lysates of HFF cells infected with TR UL5-2StrepII-2FLAG or TR (negative control - C) at an MOI of 5, prepared at the indicated times p.i. pUL5 was detected with an anti-FLAG monoclonal antibody; anti-IE1 and anti-gB monoclonal antibodies were used as markers for immediate-early and late expression kinetics, respectively, while GAPDH detection was used as a protein loading control.

The predicted molecular weight for the tagged protein is 24.5 kDa, consistent with the higher band present in the Western blot. This suggests that both forms do not undergo glycosylation. To confirm this, HEK-293T cells were transfected with pcDNA-UL5-*c-myc*-6his (calculated molecular weight 22 kDa). Lysates were treated with EndoH, PNGaseF and O-glycosidase as described in the material and methods section, subjected to SDS-PAGE and immunoblot using an anti-his tag monoclonal antibody (Fig. 5A). As expected, no band shift was observed after the treatment, demonstrating that neither form of pUL5 is glycosylated. To confirm that no glycosylation occurs even in the contest of infection, HFF cells were infected with HCMV TR-UL5-2StrepII-2FLAG at MOI 3. 72 h p.i. cell extracts were prepared and treated with the same enzymes (Fig. 5B). Similar results were obtained, confirming that UL5 protein product is not glycosylated.

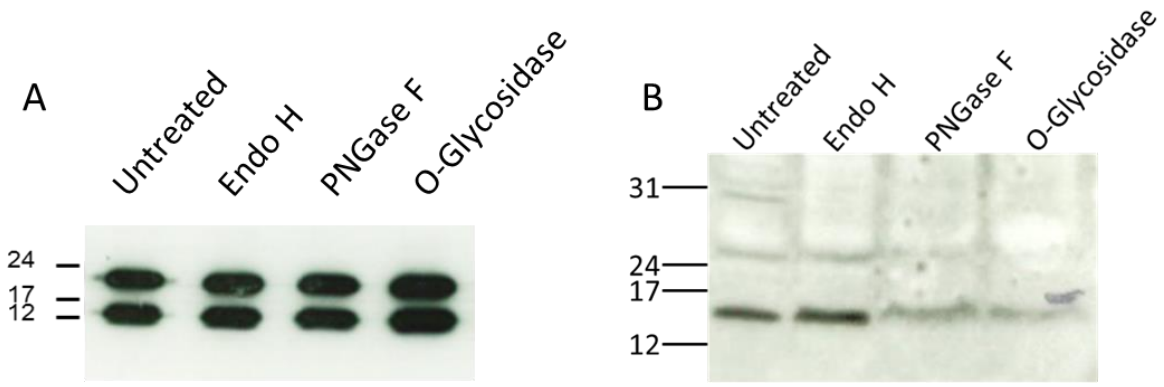

Figure 5. Glycosylation analysis of UL5 protein. Lysates from HEK-293T cells transfected with pcDNA-UL5-*c-myc*-6his (A) or from TR UL5-2StrepII-2FLAG infected HFF cells (B) were treated with Endoglycosidase H, Peptide N-glycosidase F (PNGase F) or O-glycosidase and subjected to SDS-PAGE and immunoblot with anti-his (A) or anti-Flag (B) monoclonal antibodies.

A possible explanation of the two different species of UL5 protein is the presence of a second ATG within *UL5* coding sequence. The predicted molecular weight of the amino-acid sequence from the second methionine (Met87) is 15 kDa, consistent with the lower molecular weight species detected by immunoblot both in transfection and infection environments. To test this hypothesis, a site-directed mutagenesis was carried out on the pcDNA-UL5-*c-myc*-6his, changing the ATG codon for Met87 to GCG. HEK-293T cells were transfected with pcDNA-UL5-M87A-*c-myc*-6his and lysates were subjected to SDS-PAGE and immunoblot using an anti-his tag monoclonal antibody. As seen in figure 6, the mutation M87A resulted in a strong reduction of the expression of the lower molecular weight species of UL5, suggesting that this form is translated from Met87.

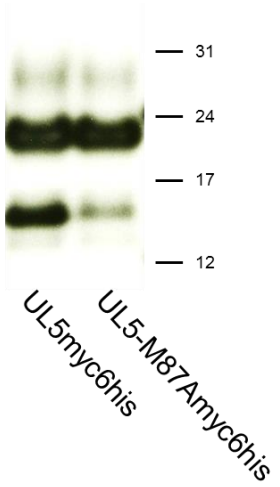

Figure 6. HEK-293T cells were tranfected with pcDNA-UL5-*c-myc*-6his or pcDNA-UL5-M87A-myc-6his and subjected to SDS-PAGE and immunoblot using an anti-his monoclonal antibody.

The addition of a combination of four tags (two StrepII and two FLAG tags) to a per-se very small protein could result in many misinterpretations. For example, the biochemical characteristics of the tagged protein as well as the subcellular localization and trafficking could be perturbed by the presence of this 54 amino-acids tail. We therefore needed a tool for the detection of wild-type UL5 protein during infection, in order to exclude the possibility of a different behavior of the protein due to the addition of the tags. To this end, we produced a recombinant his-tagged form of UL5 protein lacking the putative transmembrane region in *E. coli* BL21(DE3) strain and affinity purify it from the insoluble fraction using a NiNTA sepharose resin. The purified recombinant protein, obtained with good yield and purity as can be seen in figure 7, was formulated with Freund's adjuvant and used to immunize mice, as described in the material and methods section, to obtain an anti-UL5 serum.

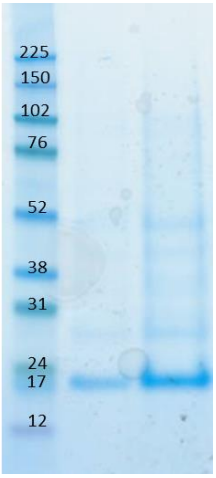

Figure 7. SDS-PAGE of 5 µg (lane 1) or 13 µg (lane 2) of purified recombinant UL5ΔTM.

LOCALIZATION OF pUL5

In order to investigate the subcellular localization of pUL5 during infection, both anti-FLAG monoclonal antibody and the antiserum raised in mice were used to immuno-detect the protein by confocal microscopy in comparison with markers of the trans-Golgi network and secretory pathway (TGN46) and glycoprotein H (gH). As displayed in Fig. 8, UL5 protein co-localizes with TGN46 and gH, suggesting that the protein accumulates in the assembly compartment (AC), where the virus particles are generated. HCMV infection changes the morphology of the nucleus which becomes larger and wraps around this juxtanuclear compartment through reorganization of cellular organelles including the endoplasmic reticulum, the Golgi apparatus, secretory vesicles and the early endosomes, as well as the lysosomes [24]. In this compartment, the nucleocapsids transported from the nucleus associate with tegument and envelope proteins with a mechanism that is not yet well understood [25]. Furthermore, we can affirm that there are no differences between the tagged (second and third rows of panels – Fig 8) and untagged (first row of panels – Fig 8) forms of pUL5, demonstrating that the addition of the tags does not affect the protein localization.

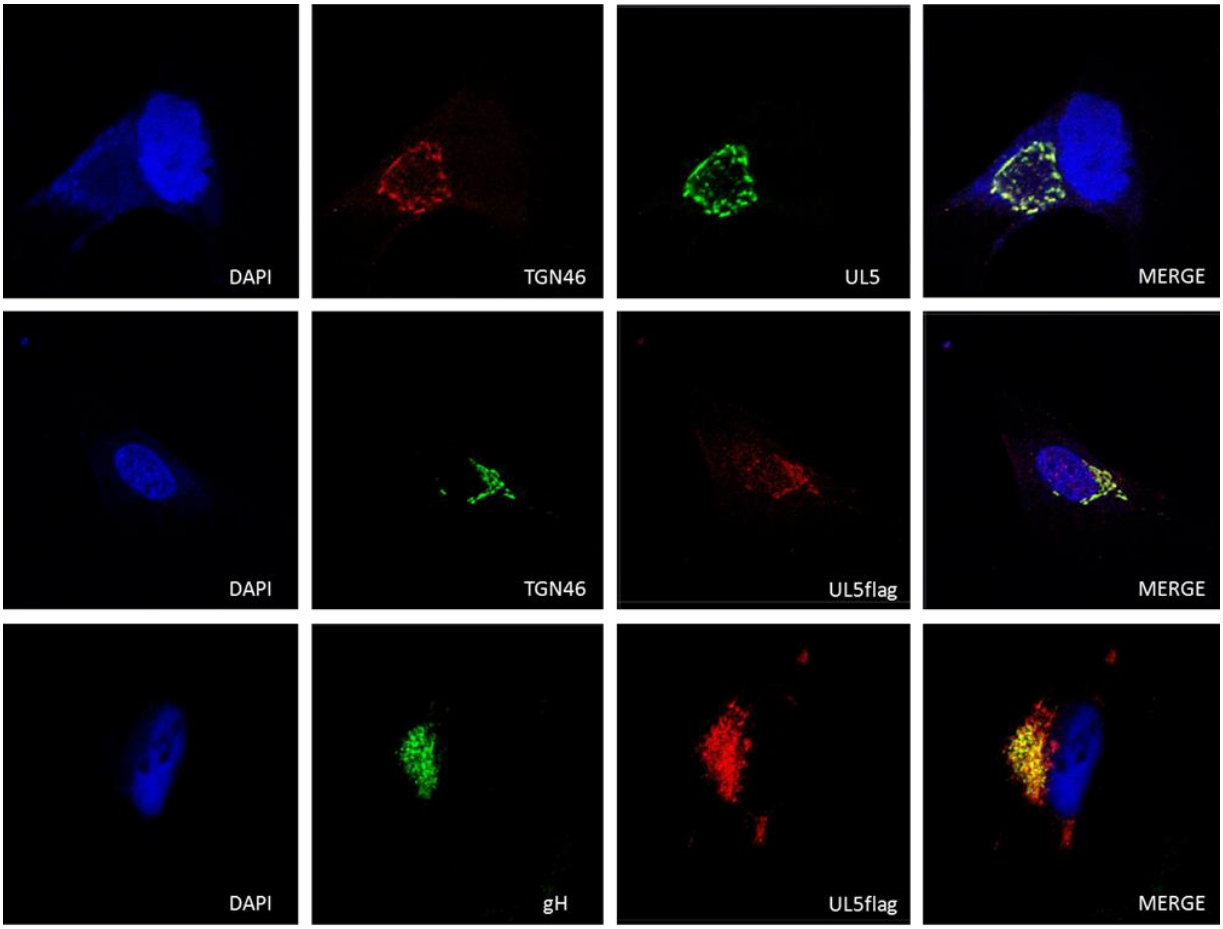

Figure 8. Subcellular localization of pUL5 in infected cells. Confocal imaging of HFF cells infected with the TR wild-type virus (first row) or with the recombinant UL5-2StrepII-2FLAG virus (second and third rows) at an MOI of 5 and fixed at 96 h p.i. The intracellular localization was determined by comparing the signal from pUL5 antiserum raised in mice or FLAG-specific antibody M2 (pUL5) with those of antibodies specific for HCMV envelope glycoprotein H and TGN46, a component of the secretory pathway trans-Golgi network. The merge panel shows the colocalization of the signals, cell nuclei are also stained blue.

Since pUL5 localizes in the AC during infection, we wondered whether it is also incorporated into the viral particles. To address this question, supernatants from TR-UL5-2StrepII-2FLAG HCMV infected fibroblasts were subjected to a sodium/tartrate gradient ultracentrifugation and mature virions were isolated and separated into the envelope and tegument fractions as described in the materials and methods section. The fractions were subjected to SDS-PAGE and immunoblot using an anti-FLAG antibody to detect the presence of UL5 protein, anti-glycopretein B and anti-pp65 antibodies were used as controls (Fig. 9).

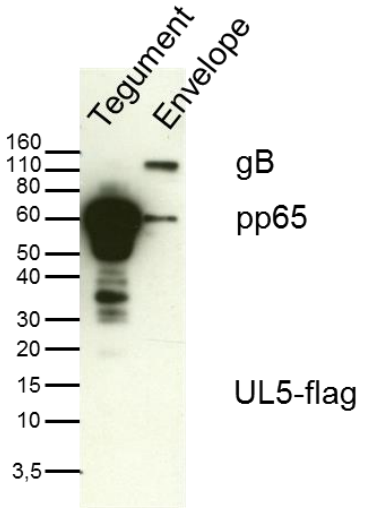

Figure 9. Western blot of envelope and tegument fractions of purified mature TR-UL5-2StrepII-2Flag HCMV virions. Anti-FLAG monoclonal antibody was used to reveal pUL5 and anti-glycoprotein B and anti-pp65 were used as controls.

As expected, glycoprotein B is only found in the envelope fraction while pp65 is almost entirely detected in the tegument fraction. In fact, pp65 is the most abundant (up to 15% of the tegument protein content) and immunogenic protein of the HCMV tegument [26]. Neither the anti-FLAG antibody (Fig. 9) nor the pUL5 anti-serum (data not shown) was able to detect UL5 protein in purified virions. Immunoblot analysis was performed both on the complete virions and dense bodies and also on separate tegument and envelope fractions and in no cases UL5 protein was identified. In 2004 Varnum et al. published a study that aimed to charachterize the HCMV virion proteome [26]. In this study, Varnum and coworkers were only able to detect a single pUL5 peptide by Fourier transform ion cyclotron resonance (FTICR) mass spectrometry in purified virions, while the less sensitive two-dimensional capillary liquid chromatography-tandem mass spectrometry (LC-MS/MS) failed to identify any pUL5 peptide. Together these data suggest that pUL5 is not a structural component of the HCMV virion and, probably, the single pUL5 peptide identified by Varnum and coworkers is the result of a random incorporation of pUL5 into the mature virion. In fact, it is very common

that several cellular and viral elements, including proteins, small molecules and RNAs, are randomly included in the virions [27].

DELETION OF *UL5* RESULTS IN INPAIRED REPLICATION IN HUMAN FIBROBLASTS

In 1991, Ripalti *et al.* published that all products of HCMV genes *UL1-UL7* are dispensable for viral growth in cell culture [28]. In order to confirm this and have an insight in pUL5 role during HCMV infection, a *UL5* deletion mutant was constructed by BAC mutagenesis as described in the materials and method part. The replication kinetics of the deletion mutant was compared with the wild-type TR HCMV to assess whether the absence of pUL5, despite being dispensable, had an effect on viral replication. Fibroblasts were infected in 12-well plates and infectious supernatants were harvested at different time points p.i. and titrated by IE-1 immuno-detection. In Fig. 10 the replication kinetics, expressed as infectious units per ml over time, of the *UL5* deletion mutant and the wild type viruses are shown.

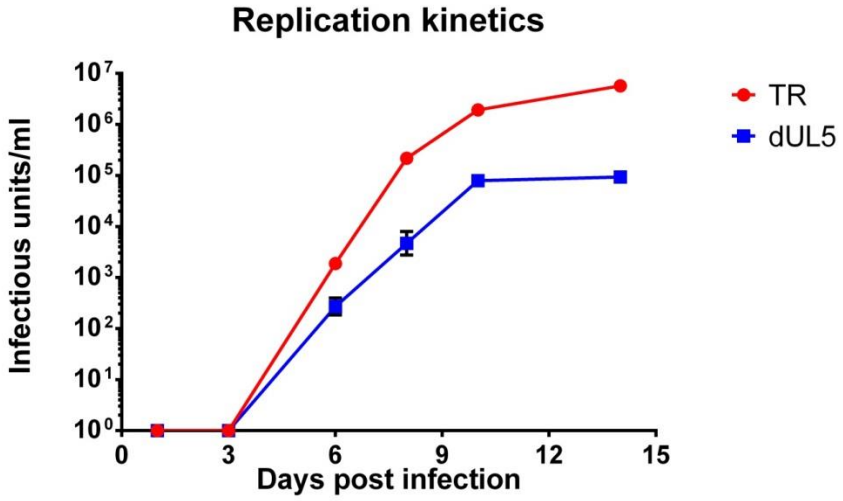

Figure 10. Replication kinetics of *UL5* deletion mutant virus with respect to the TR wild-type virus. At the indicated days p.i., supernatants from the infected cultures were harvested and titrated using indirect immunofluorescence with an antibody directed against IE1. The experiment was repeated two times, and a representative result is shown.

As Fig. 10 shows, it is evident that the absence of *UL5* ORF has a strong impact on the replication kinetics of the virus, decreasing the overall titer by two logs and showing a reduced rate constant. This may be either a direct effect or a polar effect of the mutation on the genes up and downstream *UL5* ORF, for example changing the stability of the *UL4-UL5* bicistronic

transcript or deleting sequences included in *UL5* ORF that could be involved in the regulation of other proteins expression. In order to confirm that this phenotype is a direct effect of the UL5 protein missing, further experiments complementing pUL5 are needed.

The impaired replication of *UL5* deletion mutant in contrast to the wild-type TR HCMV suggests an important role of UL5 protein in viral replication. To better understand the role of UL5 protein during HCMV infection and the reasons of why deleting *UL5* ORF results in an impaired replication, we searched for pUL5 interactors through a pull-down experiment coupled with mass spectrometry analysis. To this end, fibroblasts were infected with TR-UL5-2strepII-2FLAG and with TR wild-type (negative control) HCMV at MOI 3. Lysates were prepared at four days p.i. and subjected to immuno-precipitation (IP) using anti-FLAG monoclonal antibody-coupled magnetic beads as described in the materials and methods section. The eluted fractions were then run on SDS-PAGE and stained with colloidal Coomassie (Fig. 11A). Part of the samples was also controlled by Western blot to assess the presence of the immuno-precipitated pUL5 only in the TR-UL5-2strepII-2FLAG sample and its absence in the control using an anti-FLAG monoclonal antibody (Fig. 11B).

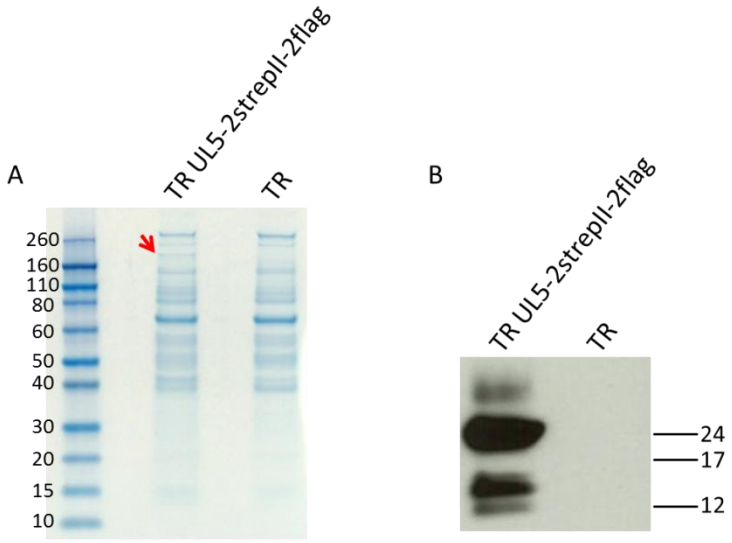

Figure 10. Pull-down of UL5-strepII-2FLAG from lysates of HFF cells infected with the recombinant virus and the control. Eluted fractions of the pull-down were subjected to SDS-PAGE and stained with colloidal coomassie (A) or revealed by immunoblot with anti-FLAG monoclonal antibody (B).

Fig. 11A shows that the pattern of the immuno-precipitation is very similar in the sample with respect to the control, with the exception of one high molecular weight species indicated by the red arrow. In order to identify it, the band of this species and the corresponding slice in the control sample were cut, digested with trypsin and analyzed by mass spectrometry. The resulting spectra were searched on Mascot database (MatrixScience) and 18 peptides of the cellular protein IQGAP1 were identified in the sample and none in the control. In order to confirm this interaction HEK-293T cells were transfected with pcDNA-UL5-*c-myc*-6his and reciprocal IPs were carried out using Dynabeads® protein G with anti-*c-myc*, anti-IQGAP1 and anti-GAPDH (negative control) antibodies as described in the materials and methods. Samples were run in SDS-PAGE and revealed by immunoblotting using the same antibodies (Fig.12).

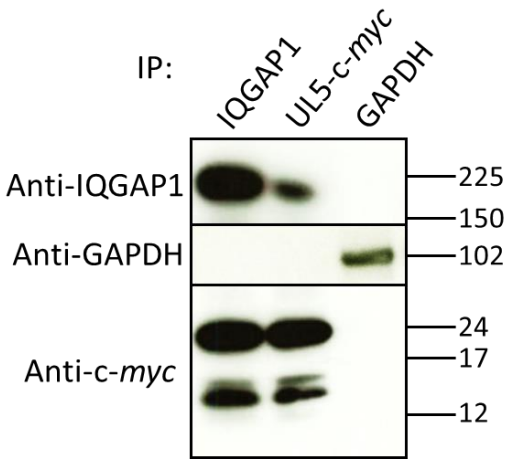

Figure 12. Immunoprecipitation of UL5 protein and IQGAP1. Lysates of HEK-293T transfected with pcDNA-UL5-*c-myc*-6his were immuno-precipitated with Dynabeads® Protein G using anti-IQGAP1, anti-*c-myc* or anti-GAPDH monoclonal antibodies. Eluted fractions were analyzed by SDS-PAGE and immunoblot with the indicated antibodies.

The IP of pUL5 with the anti-c-*myc* antibody co-precipitated IQGAP1, confirming the pull-down, and viceversa the IP with the anti-IQGAP1 antibody co-immunoprecipitated pUL5. On the other hand, IP with an unrelated antibody, anti-GAPDH, failed to precipitate both pUL5 and IQGAP1. We therefore concluded that pUL5 is a new IQGAP1 interactor.

IQGAP1 is a ubiquitously expressed scaffold protein which participates in multiple cellular processes [29], including Ca^{2+}/Calmodulin signalling, cytoskeletal organization, CDC42 and Rac signalling, E-cadherin-mediated cell–cell adhesion and β-catenin-mediated transcription [30]. It is composed of four domains: a calponin homology domain (CHD) that binds F-actin, a WW domain, a tandem repeat of four IQ motifs binding calmodulin and the GAP-related domain (GRD) mediating the binding of the Rho GTPases, CDC42 and Rac1 [30]. Despite the name, IQGAP1 lacks GTPases activating protein (GAP) activity, in fact it inhibits CDC42 GTPase activity stabilizing the receptor in its active state [31]. Overall, IQGAP1 has a key role in the modulation of actin cytoskeleton through its interaction with Rac1 and CDC42. Furthermore, IQGAP1 is deeply involved in cell–cell adhesion and migration [31]. It is, therefore, of no surprise that IQGAP1 is a common target for intracellular pathogens. For example, *Salmonella typhimurium* modulates IQGAP1 and its binding to both Rac1/Cdc42 and actin to gain entry into host cells. Furthermore, Ca2+/calmodulin signaling by enteropathogenic *E. coli* (EPEC) to induce actin pedestal formation is mediated entirely through IQGAP1 [32]. IQGAP1 has been also identified as a viral target during Ebola egress from the host cell [33]. Since this protein is involved in so many cellular mechanisms, pUL5 interaction with IQGAP1 could have several purposes. UL5 protein localizes in the AC during infection, this could suggest a role of pUL5 in the formation of this compartment by interacting with a cytoskeleton regulator. Further experiments are needed in order to understand biological and pathological meaning of this interaction.

CONCLUSIONS

We have investigated the protein product of *UL5*, a previously uncharacterized ORF of HCMV genome. *UL5* is classified as an RL11 family gene and is predicted to encode for a type Ib membrane protein with no signal peptide and 7 potential O-glycosylation sites. Northern blot analysis with a *UL5*-specific probe identified a single 1.5 kb mRNA transcribed with early kinetics, in line with what was reported in literature about a *UL4-UL5* co-transcript. By BAC mutagenesis we were able to produce recombinant viruses with a tagged version of UL5 protein allowing the identification of two forms of approximately 24 and 15 kDa expressed with late and early kinetics, respectively. We also demonstrated that neither form of UL5 protein is glycosylated and we suggest that the lower molecular weight species is the result of the translation starting from a second ATG of *UL5* ORF. Furthermore, Western blot on purified HCMV virions failed to identify UL5 protein as a structural component, suggesting that this protein is only expressed in the host cell during infection. Confocal experiments localize pUL5 with markers of the TGN and secretory pathways and with gH in the AC. We also studied the replication kinetics of the *UL5* deletion mutant virus, obtained by BAC mutagenesis, which showed a reduced replication rate in comparison to that of the wild-type virus. Finally, pull-down experiments coupled with mass spectrometry analysis identified the cellular scaffold protein IQGAP1 as a UL5 protein interactor and reciprocal IPs in transfection further confirmed this interaction. Further studies are needed to understand the function of pUL5 during HCMV infection.

Cells lines and antibodies. Human foreskin fibroblast (HFF), human fetal lung fibroblast (MRC-5) and human embryonic kidney 293 (HEK-293T) cell lines were purchased from ATCC (catalogue numbers SCRC-1041, CCL-171 and CRL-11268 respectively) and cultured according to supplier's instructions. Dulbecco's modified Eagle's medium (DMEM) high glucose (Gibco, Life Technologies) was supplemented with 10% fetal calf serum (FCS) and penicillin streptomycin glutamine (Gibco, Life Technologies).

Primary antibodies used were: mouse anti-Flag (Sigma-Aldrich), mouse anti-6His (Life Technologies), mouse anti-*c-myc* (Life Technologies), rabbit anti-FLAG (Sigma-Aldrich), mouse anti-pp28 (Vyruses), mouse anti-gH (AbD Serotech), sheep anti-TGN46 (AbD Serotech), mouse anti-PDI (Life Technologies), rabbit anti-PDI (Abcam), mouse anti-EEA1(Abcam), rabbit anti-EEA1 (Abcam), mouse anti-gB (Abcam) and mouse anti-GM130 (Abcam) monoclonal antibodies. Secondary antibodies used were: Alexa Fluor $F(ab)_2$ fragment of 488-, 568-, 647- conjugated goat anti-mouse/anti-rabbit (Life Technologies) and horseradish peroxidase-conjugated secondary antibodies from Perkin Elmer. Wheat Germ Agglutinin (WGA) 647-conjugated was purchased from Life Technologies.

Viruses. Wild-type (WT) HCMV (TR strain) was derived from a bacterial artificial chromosome (BAC) containing the HCMV TR genome. The BAC TR was used to create recombinant viral genomes via a marker-less two-step RED-GAM BAC mutagenesis [22]. Briefly, kanamycin resistance cassette, flanked by I-SceI restriction enzyme cleavage sites, was amplified from pEP-TAP shuttle vector using primers containing homologous regions for the integration in the region of interest. All the primers used are listed in Table 1. Recombination events were performed with *E. coli* GS1783 strain containing a BAC clone of the HCMV TR strain, the lambda Red system under the control of a heat-inducible promoter and the I-SceI genes under the control of an arabinose-inducible promoter. The first recombination step consists in the electroporation of the purified PCR-amplified cassette in

competent, heat-induced GS1783 cells. Positive clones for cassette integration were selected based on kanamycin resistance and screened both by PCR and sequencing. The second recombination was triggered through both heat-shock and arabinose and results in the excision of the kanamycin resistance, leaving the tag sequence (V5-Streptavidin binding peptide or 2xstrepII-2xFLAG) in frame with the C-terminus of the gene of interest. Presumptive clones were screened by PCR and sequencing, while the integrity of the recombinant HCMV genome was assessed through HindIII and BamHI restriction analysis. To reconstitute the virus, actively growing MRC-5 cells from a nearly-confluent T175-flask were trypsin detached, mixed with about 3 μg fresh prepared BAC DNA and 1 μg pCMVKm2-pp71 plasmid and electroporated in 4 mm cuvettes at 250 V and 950 μF. Supernatant was collected from infected cells when cytopathic effect was >90%. For all following infections, HFF cells were used.

***UL5* RNA analysis.** For Northern blot analysis, whole-cell RNA was isolated from mock-infected or infected HFF cells at 6, 24, 48, and 72 h postinfection (p.i.) using RNeasy minikit (Qiagen). 10 μg of each sample were separated on 1% agarose gels and Northern blot analysis was carried out using NorthernMax® kit (Life Technologies). 500 ng of each sample was also run on a separate gel as a control. Hybridization and detection were performed by using the BrightStar® BioDetect™ kit (Life Technologies) according to the manufacturer's instructions. Briefly, *UL5*-specific DNA hybridization probes were labeled with biotin-16-dUTP (Roche) by PCR with *UL5*-specific primers UL5-NB-F and UL5-NB-R (Table 1), using TR BAC as a template. The PCR mix was prepared with 40% of biotinylated dUTP and 60% of dTTP.

Preparation of eukaryotic expression constructs and DNA transfection. Lipofectamine 2000 (Life Technologies) was used to transfect HEK-293T cells according to manufacturer's instructions. Human codon-optimized *UL5* gene, from HCMV TR strain, was synthesized by Geneart (Life Technologies) and cloned in plasmid pcDNA3.1(-)/*c-myc*/6HisA (Life Technologies) in frame with C-term *c-myc* and six histidine tag sequences. pUL5 soluble

form was obtained by deleting the putative trans-membrane region (nucleotides 319 to 498) from the pcDNA3.1(-)UL5/*c-myc*/6HisA using Quickchange® site directed mutagenesis (Stratagene) according to manufacturer's instructions with primers UL5ΔTM-F and UL5ΔTM-R (Table 1). Correct deletion was confirmed by DNA sequencing. Mutation of the second ATG in *UL5* open-reading frame (Met87) was obtained with the same strategy using primers UL5-M87A-F and UL5-M87A-R (Table 1).

Production of recombinant his-tagged proteins in *E. coli* and antisera. All constructs for recombinant protein expression in *E. coli* were cloned using the Polymerase Incomplete Primer Extension (PIPE) method [34] with KAPA HiFi HotStart ReadyMix (kapabiosystems). Insert PCR for UL5ΔTM was amplified from pcDNA3.1(-)UL5ΔTM/*c-myc*/6HisA with primers UL5Ec F and UL5Ec R, listed in Table 1, treated with DpnI (New England Biolabs - NEB) to eliminate the template and purified using Wizard® SV Gel and PCR Clean-Up System. Vector PCR of pET15Tev was amplified with the primers petTEVfor and petTEVrev. Vector and DpnI-treated Insert PCRs were transformed in Mach1™ competent cells (Life Technologies). Positive clones were screened by colony PCR and sequencing. About 150 ng of plasmids were transformed in BL21(DE3) competent cells for recombinant protein expression using EnPresso™ B system (Biosilta) according to manufacturer's instructions. The lysis was performed by sonication and the recombinant proteins were purified from inclusion bodies on a Ni sepharose HisTrap column (GE Healthcare). pUL5 was solubilized in Urea 8M and purified in denaturing conditions. Briefly, the resin was equilibrated with 10 column volumes (CV) of Binding buffer (20 mM Tris-HCl, 300 mM NaCl, 10 mM Imidazole, 8 M Urea, 5 mM di-tio-tritole pH 8.0) before loading the sample, after binding step 20 CV of Binding buffer were used for washing step. Protein was eluted in 1 mL fractions with Elution buffer (20 mM Tris-HCl, 300 mM NaCl, 300 mM Imidazole, 8 M Urea, 5 mM di-tio-tritole pH 8.0). Protein purity was evaluated by SDS-PAGE. The protein concentration was measured with BCA protein assay (Thermo Scientific). Recombinant purified pUL5 was injected in a group of 4 mice with 3 doses, once per 14 days,

of 20 μg of recombinant protein formulated in Freund adjuvant. At the end of the protocol the mice were sacrificed and sera collected.

Confocal microscopy analysis. HFF cells grown on glass cover slips in 24-well plates were infected at MOI 5. At different time points after infection, the cover slips were washed and fixed in Cytofix/Cytoperm™ buffer (BD Biosciences). The fixed cells were incubated in blocking buffer (phosphate-buffered saline - PBS + 5% non-immune human serum) for 30 minutes before antibody staining. Primary and secondary antibodies were always diluted in blocking buffer and incubated for 30 min at room temperature. Coverslips were mounted using ProLong® Gold Antifade Mountant with DAPI (Life Technologies). The intracellular staining was examined under laser illumination in a Zeiss LMS 710 confocal microscope and images were captured using ZEN software (Carl Zeiss). Images were processed using ImageJ software.

SDS-PAGE and immunoblotting. Proteins were separated by sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on 4-12% polyacrylamide pre-cast gels (Life Technologies) under standard conditions. Proteins were transferred to nitrocellulose membranes (iBlot® system – Life Technologies), and membranes were blocked with PBS containing 0.05% Tween 20 and 10% powdered milk. Antibodies and sera were diluted in PBS containing 0.05% Tween 20 and 1% powdered milk. For detection of primary antibody binding, horseradish peroxidase-conjugated anti-rabbit or anti-mouse IgG antibody and the West Pico Chemoluminescent Substrate (Thermo Scientific) were used according to the manufacturer's instructions. Removal of N-linked oligosaccharides was carried out using recombinant peptide:N-glycosidase F (PNGase F) and endoglycosidase H (Endo H) (NEB) according to the manufacturer's specification.

HCMV Virions purification and virion protein fractionation. Mature HCMV virions were separated from non-infectious enveloped particles (NIEPs) and dense bodies (DB) through a positive density/negative viscosity step-gradient centrifugation as described previously [35]. Briefly, supernatant of infected cells was collected when 100% of cytopathic

effect was observed and subjected to 4000 rpm centrifugation for 20 minutes at 4°C. Clear supernatant was transferred to polycarbonate ultracentrifuge tubes under lied with 20% sucrose cushion and centrifuged at 23000 rpm in a Beckman SW32Ti rotor for 60 minutes. The pelleted virus was resuspended in 1 ml PBS, 2.5 ml of 4 different solutions containing decreasing concentration of sucrose and increasing concentration of glycerol tartrate were layered underneath. Tubes were centrifuged at 42000 rpm for 60 min at 10°C in a Beckman SW28Ti rotor. Bands containing mature virions were collected through a syringe 25 gauge, resuspended in PBS and centrifuged at 23000 rpm in a Beckman SW32Ti rotor for 60 min. The virus-containing pellet was resuspended in PBS. To separate envelope from capsid and tegument proteins, purified virions were mixed 1:1 with envelope extraction buffer (1% Nonidet-P (NP)-40 and 4% Sodium Deoxycholate) and incubated on ice for 30 min with occasional vortexing. Soluble fraction (envelope) was collected through centrifugation at 13000 rpm in a benchtop centrifuge for 30 min at 4°C. The insoluble pellet (capsid and tegument) was washed twice with PBS before being solubilized in SDS-PAGE sample buffer. For each extraction, a total of 4 confluent T175 cm^2 flasks of HFF cells were infected.

Viral replication kinetics and titration. To assay the replication kinetics of the viral mutants, one day after seeding $1.25x10^5$ HFF cells per well in 12-well plates, cells were infected in quadruplicates with 500µl of different dilutions of infectious supernatant at 37°C for 3 h. After incubation, the infectious inoculum was removed and fresh medium was added. The next day, one of the four cultures of each dilution was assayed to determine the percentage of infection via immediate early antigen 1 immuno-staining. The dilutions resulting in approximately 60% of infected cells were chosen to assay the viral replication kinetics. To this end, supernatant from the three remaining cultures was taken at different time points after infection and stored at room temperature until titration. To determine the virus titer, $3x10^5$ HFF cells per well were seeded in a 96-well plate the day before the experiment. Virus stocks were serially diluted l0fold and 50 µl of diluted virus was seeded in triplicate wells. After a 3 h incubation at 37°C, the virus containing medium was removed and replaced with fresh

medium. After 16 h incubation the medium was removed and the plate washed trice with Dulbecco's phosphate buffered saline (DPBS, pH 7.4) and fixed in absolute ethanol for 15 min at room temperature. Following removal of the ethanol, the plate was rehydrated with DPBS for 20 min at room temperature. The wells were then incubated with 0.1-0.3 µg per well of mouse anti-IE1 diluted in DPBS for 1 h at 37°C. After washing the wells five times with DPBS, 0.1 ml of FITC-conjugated goat anti-mouse IgG diluted l/100 in DPBS was added and the plate incubated for 1 h at 37°C. The wells were then washed twice with DPBS and 50µl of DPBS containing 30% glycerol was added. The plate was viewed with Zeiss Axiovert 200 fluorescence microscope with a stage adapter for reading multiwell plates. Fluorescent nuclei were counted and infectivity expressed as infectious units IU (fluorescent nuclei) per ml of inoculum.

Pull-down coupled with mass spectrometry and immunoprecipitations. For pull-down experiments HFF cells were infected with the recombinant or wild-type viruses at MOI 3. Cell lysates were prepared with CelLytic M (Sigma-Aldrich) according to manufacturer's specifications. Pull-down was carried out using anti-FLAG® M2 magnetic beads (Sigma-Aldrich) according to manufacturer's instructions; the binding step was prolonged overnight and protein complexes were eluted by competition with the FLAG peptide. Eluted fractions were subjected to SDS-PAGE and stained with colloidal Coomassie G-250. Bands of interest were cut, dehydrated with acetonitrile for 15 minutes and digested with trypsin overnight at 37°C, then spotted on the target 1:1 with recrystallization buffer (70% ethanol, 1% trifluoroacetic acid) and analyzed by MALDI-TOF. The trypsin spectrum was subtracted from the resulting spectra and the final peaks were searched on Mascot database (MatrixScience). For immunoprecipitations (IP), HEK-293T cells were transfected with Lipofectamine 2000 (Life Technologies) and lysates were prepared using CelLytic M (Sigma-Aldrich). IPs were carried out using Dynabeads® Protein G (Life Technologies) according to manufacturer's specifications, using 5 µg of antibody for IP reaction.

Sequence analysis. The amino-acid sequences for the UL5 coding sequences from laboratory adapted HCMV strains and clinical isolates have been assigned the following accession numbers from GenBank: AGL96607.1, AAS48922.1, AFR56339.1, AFR56172.1, AFR56005.1, AFR54510.1, AGT36439.1, ACZ80259.1, ACZ80094.1, ACZ79929.1, ACZ79764.1, ACS92108.1, ACM47996.1, ACL51086.1, AAR31292.1 and AAR31275.1. Multiple alignment was performed using CLUSTALW [14]; N-, and O-glycosylation sites, signal peptides and transmembrane regions were predicted using NetNGlyc [17], NetOGlyc [18], SignalP [16] and PSORTII [15] servers, respectively.

Table 1: List of oligos

Name	Sequence
UL5-SBP-V5 F	GCCCGTCCTTTGAAATACCGCTATCAACGTCTCGTCGCCACCGTGGGCAAGCCCATCCCCAACCCCCTG
UL5-SBP-V5 R	CAAATACGCAAAAGCAAAACACAACAAACTATACACAGCTGGCTAACTAGCTAGGGCTCTCTCTGGCCCTGGGGGTGGTGCTCCAGTCTGGCCAACCAATTAACCAATTCTGATTAG
UL5-2flag2strep F	GCCCGTCCTTTGAAATACCGCTATCAACGTCTCGTCGCCACCGTGTGGAGCCACCCGCAGTTCGAAAAGGGCGGCGGC
UL5-2flag2strep R	CAAATACGCAAAAGCAAAACACAACAAACTATACACAGCTGGCTAACTAGCTACTTATCGTCGTCATCCTTGTAATCGCTGCCGCCGCCGCCCCGATTTATTCAACAAAGCCACGTTG
ΔUL5 F	TGTTTACGTTGCTTTTGAAATGTTAAGCGTCCCTACGGCGCTAACTAGGGATAACAGGGTAATCGATTTATTC
ΔUL5 R	CAAATACGCAAAAGCAAAACACAACAAACTATACACAGCTGGCTAACTAGGTTAGCGCCGTAGGGACGCTTAACATTTCAAAAGCAACGGCCAGTGTTACAACCAATTAACCAATTC
UL5Ec F	CTGTACTTCCAGGGCATGTTTCTGGGCTACAGCGACTG
UL5Ec R	AATTAAGTCGCGTTAGCCGCCGTAGAACTGCTTGGTGATG
UL5dTM-F	GATCATCACCAAGCAGGGATCCGAGCTCGGTACC
UL5dTM-R	GGTACCGAGCTCGGATCCCTGCTTGGTGATGATC
HindT7F	TTATAAGCTTATGAACACGATTAACATCGCTAAG
T7NheR	ATATGCTAGCTTACGCGAACGCGAAGTCCGACTCTAAG
CMVp-dNcoI-F	CTATTACCTTGGTGATGCGGTTTTGGCAG
CMVp-dNcoI-R	TCACCAAGGTAATAGCGATGACTAATAC
petTEVfor	TAACGCGACTTAATTCTAGCATAACCCCTTGGGGCCTCAAACGG
petTEVrev	GCCCTGGAAGTACAGGTTTTCGTGATGATGATGATGATGGCTGCTGCCCATGGTATATC
UL5-M87A-F	CCCGATAGAAGTGCGCTGCTGTCCCGGGAAGAGG
UL5-M87A-R	CGCACTTCTATCGGGAGGGGGGGCAATG
UL5-NB-F	CTCTGACTGTGTAGATCCCGGC
UL5-NB-R	CTGTCGTTTCTGCATAGACCGTCTC

REFERENCES

1. ten Berge, I.J.M. and R.A.W. van Lier, *The interaction between cytomegalovirus and the human immune system.* Immunology Letters, 2014. **162**(2, Part B): p. 141-144.

2. Davison, A.J., et al., The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. Journal of General Virology, 2003. **84**(1): p. 17-28.

3. Davison, A.J., et al., Homology between the human cytomegalovirus RL11 gene family and human adenovirus E3 genes. Journal of General Virology, 2003. **84**(3): p. 657-663.

4. Chee, M.S., et al., Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr Top Microbiol Immunol, 1990. **154**: p. 125-69.

5. Atalay, R., et al., Identification and Expression of Human Cytomegalovirus Transcription Units Coding for Two Distinct Fcγ Receptor Homologs. Journal of Virology, 2002. **76**(17): p. 8596-8608.

6. Cortese, M., et al., Recombinant Human Cytomegalovirus (HCMV) RL13 Binds Human Immunoglobulin G Fc. PLoS ONE, 2012. **7**(11): p. e50166.

7. Stanton, R.J., et al., Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication. The Journal of Clinical Investigation, 2010. **120**(9): p. 3191-3208.

8. Chang, C.P., et al., Identification and expression of a human cytomegalovirus early glycoprotein. Journal of Virology, 1989. **63**(8): p. 3330-3337.

9. Shikhagaie, M., et al., The Human Cytomegalovirus-Specific UL1 Gene Encodes a Late-Phase Glycoprotein Incorporated in the Virion Envelope. Journal of Virology, 2012. **86**(8): p. 4091-4101.

10. Gabaev, I., et al., Expression of the Human Cytomegalovirus UL11 Glycoprotein in Viral Infection and Evaluation of Its Effect on Virus-Specific CD8 T Cells. Journal of Virology, 2014. **88**(24): p. 14326-14339.

11. Gabaev, I., et al., The Human Cytomegalovirus UL11 Protein Interacts with the Receptor Tyrosine Phosphatase CD45, Resulting in Functional Paralysis of T Cells. PLoS Pathog, 2011. **7**(12): p. e1002432.

12. MacManiman, J.D., et al., Human Cytomegalovirus-Encoded pUL7 Is a Novel CEACAM1-Like Molecule Responsible for Promotion of Angiogenesis. mBio, 2014. **5**(6).

13. Engel, P., et al., Human cytomegalovirus UL7, a homologue of the SLAM-family receptor CD229, impairs cytokine production. Immunol Cell Biol, 2011. **89**(7): p. 753-766.

14. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics (Oxford, England), 2007. **23**(21): p. 2947-2948.

15. Nakai, K. and P. Horton, PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci, 1999. **24**(1): p. 34-6.

16. Petersen, T.N., et al., SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods, 2011. **8**(10): p. 785-6.

17. R., G., J. E., and B. S., Prediction of N-glycosylation sites in human proteins. 2004.

18. Steentoft, C., et al., Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. Embo j, 2013. **32**(10): p. 1478-88.

19. Chang, C.P., C.L. Malone, and M.F. Stinski, A human cytomegalovirus early gene has three inducible promoters that are regulated differentially at various times after infection. Journal of Virology, 1989. **63**(1): p. 281-290.

20. Alderete, J.P., S.J. Child, and A.P. Geballe, Abundant Early Expression of gpUL4 from a Human Cytomegalovirus Mutant Lacking a Repressive Upstream Open Reading Frame. Journal of Virology, 2001. **75**(15): p. 7188-7192.

21. Alderete, J.P., S. Jarrahian, and A.P. Geballe, Translational Effects of Mutations and Polymorphisms in a Repressive Upstream Open Reading Frame of the Human Cytomegalovirus UL4 Gene. Journal of Virology, 1999. **73**(10): p. 8330-8337.

22. Tischer, B.K., et al., Two-step Red-mediated recombination for versatile high-efficiency markerless DNA manipulation in Escherichia coli. Biotechniques, 2006. **40**: p. 191-197.

23. Li, Y., Commonly used tag combinations for tandem affinity purification. Biotechnol Appl Biochem, 2010. **55**(2): p. 73-83.

24. Alwine, J.C., The Human Cytomegalovirus Assembly Compartment: A Masterpiece of Viral Manipulation of Cellular Processes That Facilitates Assembly and Egress. PLoS Pathog, 2012. **8**(9): p. e1002878.

25. Britt, B., *Maturation and egress*, in *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, A. Arvin, et al., Editors. 2007, Cambridge University Press

Copyright (c) Cambridge University Press 2007: Cambridge.

26. Varnum, S.M., et al., Identification of Proteins in Human Cytomegalovirus (HCMV) Particles: the HCMV Proteome. Journal of Virology, 2004. **78**(20): p. 10960-10966.

27. Mocarski, E.S., T. Shenk, and R.F. Pass, Cytomegaloviruses. In: Knipe DM, Howley PM, Fields Virology. Wolters Kluwer, Lippincott Williams & Wilkins. 2013.

28. Ripalti, A. and E.S. Mocarski, *The products of human cytomegalovirus genes UL1-UL7, including gp48, are dispensable for growth in cell-culture*. Progress in Cytomegalovirus Research, ed. M.P. Landini. Vol. 978. 1991. 57-60.

29. White, C.D., H.H. Erdemir, and D.B. Sacks, *IQGAP1 and its binding proteins control diverse biological functions.* Cellular Signalling, 2012. **24**(4): p. 826-834.

30. Briggs, M.W. and D.B. Sacks, *IQGAP proteins are integral components of cytoskeletal regulation.* EMBO Reports, 2003. **4**(6): p. 571-574.

31. Noritake, J., et al., *IQGAP1: a key regulator of adhesion and migration.* Journal of Cell Science, 2005. **118**(10): p. 2085-2092.

32. Kim, H., C.D. White, and D.B. Sacks, *IQGAP1 in microbial pathogenesis: targeting the actin cytoskeleton.* FEBS letters, 2011. **585**(5): p. 723-729.

33. Lu, J., et al., Host IQGAP1 and Ebola Virus VP40 Interactions Facilitate Virus-Like Particle Egress. Journal of Virology, 2013. **87**(13): p. 7777-7780.

34. Klock, H. and S. Lesley, The Polymerase Incomplete Primer Extension (PIPE) Method Applied to High-Throughput Cloning and Site-Directed Mutagenesis, in High Throughput Protein Expression and Purification, S. Doyle, Editor. 2009, Humana Press. p. 91-103.

35. Talbot, P. and J.D. Almeida, *Human Cytomegalovirus: Purification of Enveloped Virions and Dense Bodies.* Journal of General Virology, 1977. **36**(2): p. 345-349.