frontiers
in Microbiology

HCMV Envelope Glycoprotein Diversity Demystified

Mathilde Foglierini[1,2], Jessica Marcandalli[1] and Laurent Perez[1]*

[1] Faculty of Biomedical Sciences, Institute for Research in Biomedicine, Università della Svizzera Italiana, Bellinzona, Switzerland, [2] Swiss Institute of Bioinformatics, Lausanne, Switzerland

Human cytomegalovirus (HCMV) is the leading viral cause of congenital birth defects and is responsible for morbidity and mortality in immunosuppressed individuals. Considerable efforts have been deployed over the last decade to develop a vaccine capable of preventing HCMV infection. However, in recent clinical trials, vaccines showed at best modest efficacy in preventing infection. These findings might be explained by the high level of sequence polymorphism at the genomic level. To investigate if genomic variation also leads to antigenic variation, we performed a bioinformatic sequence analysis and evaluated the percentage of conservation at the amino acid level of all the proteins present in the virion envelope. Using more than two hundred sequences per envelope glycoprotein and analyzing their degree of conservation, we observe that antigenic variation is in large part limited to three proteins. In addition, we demonstrate that the two leading vaccine candidates, the pentamer and gB complexes, are well conserved at the amino acid level. These results suggest that despite genomic polymorphism, antigenic variability is not involved in the modest efficacy observed in the recent clinical trials for a HCMV vaccine. We therefore propose that next-generation vaccines should focus on stabilizing and refining the gB domains needed to induce a protective humoral response.

Keywords: human cytomegalovirus, envelope glycoproteins, viral diversity, multiple sequence alignment, protein sequence analysis, phylogenic analysis

INTRODUCTION

Human cytomegalovirus (HCMV) is a member of the *Betaherpesvirinae* subfamily of *Herpesviridae* and 40–80% of the human adult population is seropositive to HCMV (Kenneson and Cannon, 2007). Primary infection is generally asymptomatic in immunocompetent individuals. Nonetheless, HCMV establishes lifelong latency in infected individuals (Zhuravskaya et al., 1997), and viral reactivation in immunosuppressed hosts such as transplant patients and patients with AIDS can cause severe diseases or lead even to death (Rando et al., 1990). HCMV is also the major cause of congenital birth defects (0.7% of the newborns) (Kirby, 2016), exceeding fetal alcohol syndrome and Zika virus infection in industrialized countries. Current antiviral therapies and treatments with hyperimmune globulin transfusion to control viremia are not efficient (Tandon and Mocarski, 2012). Due to the prevalence, severity and importance of this virus, the United States National Academy of Medicine considers that obtaining an effective vaccine against HCMV is a top priority in public health. However, HCMV is a complex virus with a genome encoding at minimum 170 open reading frames (ORFs) (Stern-Ginossar et al., 2012) and expressing at least nineteen

membrane proteins integral to the virion envelope (Weekes et al., 2014). These envelope proteins are primarily required for binding to the host cells, viral entry and, in some cases, immune evasion by sequestering human chemokines (Frascaroli et al., 2006; Frascaroli et al., 2018). HCMV broad cellular tropism (epithelial/endothelial cells, fibroblasts, monocytes/macrophages, smooth muscle cells, neurons, stromal cells and hepatocytes) is due to two glycoprotein complexes responsible for the interaction with specific host cell receptors (Malito et al., 2018; Nguyen and Kamil, 2018). The gHgLgO (trimer) complex is required for infection of all cell types and is known to bind the platelet-derived growth factor receptor α (PDGFRα) expressed on fibroblasts (Kabanova et al., 2016). The gHgLpUL128pUL130pUL131A (pentamer) complex is required for viral entry in epithelial, endothelial and myeloid cells and binds Neuropilin2 (Nrp2) (Martinez-Martin et al., 2018). Membrane fusion between the virus and the surface of the host cell is catalyzed by the glycoprotein B (gB), a class III fusion protein forming a homotrimer (Burke and Heldwein, 2015; Chandramouli et al., 2015). Promising HCMV vaccine candidates have been sought among the viral proteins present in the envelope and in the tegument layer of the virus (Anderholm et al., 2016). Current vaccine research is focusing on two protein complexes of the virion envelope: the gB homotrimer and the pentamer (Chiuppesi et al., 2017; Wussow et al., 2018). Both complexes were identified as dominant targets of the humoral immune response upon infection (Macagno et al., 2010; Kabanova et al., 2014). Research studies on gB are more advanced compared to the pentamer, and a vaccine was already tested in phase I and II clinical trials (Baraniak et al., 2018a; Nelson et al., 2018b; Schleiss, 2018). The vaccine demonstrated around 50% efficacy in preventing HCMV acquisition in seronegative women (Baraniak et al., 2018a; Nelson et al., 2018a). However, antibodies elicited by the vaccine were only poorly neutralizing (Baraniak et al., 2018a; Nelson et al., 2018b). To explain these results, several hypotheses have been proposed: the presence of an undesired antigenic domain (AD) in the gB antigen used in the vaccine, the post-fusion conformation of the antigen injected, but also the sequence diversity observed among the gB proteins expressed by the different natural circulating HCMV strains (Renzette et al., 2011, 2017). Indeed, HCMV is described as being a highly polymorphic virus at the genomic level (Renzette et al., 2016), and it was shown that sequence diversity of some proteins impacts the infection outcome (Renzette et al., 2011; Galitska et al., 2018). To investigate if HCMV genomic diversity is also leading to antigenic variation of the proteins belonging to the virion envelope, we used datasets from publicly available HCMV genomes and performed sequence diversity analysis.

MATERIALS AND METHODS

Sequences Analysis of HCMV Proteins

Eighteen envelope and membrane protein sequences of HCMV were obtained from the National Center for Biotechnology Information (NCBI) Entrez Protein database as of September 18, 2018 (Sayers et al., 2019) and were derived from the following genes: RL10, UL1, UL4, UL33, UL55, UL73, UL74, UL75, UL78, UL100, UL115, UL116, UL128, UL130, UL131A, UL132, US27, and US28. For each protein, the sequences were purged of duplicates for a given strain, partial sequences (+/− 5 amino acids) and sequences lacking the name of the strain, the country where it was isolated and/or the collection date. For downstream analysis, we used on average 240 sequences for each HCVM protein. Protein sequences were aligned with Clustal Ω (Sievers et al., 2011). For a given protein, a matrix presenting the percentage of amino acids identity between strains was generated followed by hierarchical clustering using the Heatmap function from the "ComplexHeatmap" R package (Gu et al., 2016) (Euclidean metric and complete aggregation method). The overall mean distances (percentage of mean identity) were computed with MEGA X (Kumar et al., 2018) with default parameters. Analyses of the glycoprotein pentamer complex gHgLpUL128pUL130pUL131A (UL75-UL115-UL128-UL130-UL131A) and gB (UL55) were conducted by concatenation of each glycoprotein for a given strain. We used the sequences of 207 strains containing the pentamer and gB to perform an evolutionary analysis. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (Jones et al., 1992), bootstrapped 500 times. The trees with the highest log likelihood are shown and are drawn to scale, with branch lengths measured in the number of substitutions per site. Evolutionary analyses were conducted in MEGA X. Trees were visualized with iTOL (Letunic and Bork, 2016). Histograms of consensus sequence for the "hot-spots" of pentamer, gB, gN, and gO were obtained by Jalview software (Waterhouse et al., 2009). The list of HCMV strains used in this study is available as (**Supplementary Table S1**).

RESULTS

Envelope Protein Required for Viral Tethering Are Well Conserved

Human cytomegalovirus envelope contains four viral G protein-coupled receptors (vGPCRs), namely pUL33, pUL78, pUS27, and pUS28 proteins (de Munnik et al., 2015). Among these four vGPCRs, only pUS28 is functional and possesses structural similarity to the human chemokine receptors CCR1 and CCR5 (Gao and Murphy, 1994). The pUS28 also binds human CX3CL1 and additional inflammatory chemokines from the CC-family (Krishna et al., 2018). Analysis by similarity matrix demonstrated that all vGPCRs are very well conserved (**Figure 1A**), a result consistent with the conserved architectural structure of the GPCRs (Congreve et al., 2017). The vGPCRs have a limited immunogenicity, due to the minimal exposure of the short loops that link the seven-(pass)-transmembrane domains. The glycoprotein M (UL100) together with glycoprotein N (UL73) forms the gM/gN heterodimer, which is one of the most abundant protein complexes within the HCMV envelope. Furthermore, this heterodimer is one of the few envelope proteins conserved among herpesviruses, thus suggesting an important function in their biology. Although the latter is not fully understood, the

complex was shown to be involved in the initial interaction phases with host cells by binding heparin present on the cellular envelope (Compton et al., 1993). Interestingly, whereas gM is extremely conserved with 99% mean identity across strains, gN is one of the less conserved proteins with only 81% of identity (**Figure 1A**). Moreover, gN is extensively modified by O-linked sugars which contribute to its mass for more than 40 kDa (Pignatelli et al., 2001). Nevertheless, gN was identified as a target of neutralizing antibodies (Shimamura et al., 2006), a surprising result considering its sequence variability and glycosylation status (**Supplementary Figure S1a**).

Structural Components of the Virion Envelope Exhibit a Low Degree of Conservation

The organization of the viral proteins in the envelope is still poorly characterized and understood. Most of the proteins from this group display a very low degree of conservation mean identity across strains (**Figure 1B**). The gp1 (UL1) is a glycoprotein with 224 amino acids and whose function remains unclear. However, it was shown that gp1 contains an Ig-like domain, and disruption of gp1 in HCMV generates viral growth defects in epithelial cells but not in fibroblasts (Shikhagaie et al., 2012). The gp48 (UL4) is a 48 kDa glycoprotein whose expression is non-essential for HCMV replication in cell cultures (Alderete et al., 2001). The gp42 (UL132) is a highly glycosylated protein of 270 amino acids, described to be required for cellular endocytosis of the virus. Furthermore, gp42 is necessary for optimal replication of the virus (Kropff et al., 2010). The gpTRL10 (RL10) is a 170-amino acid-long glycoprotein with no function yet identified.

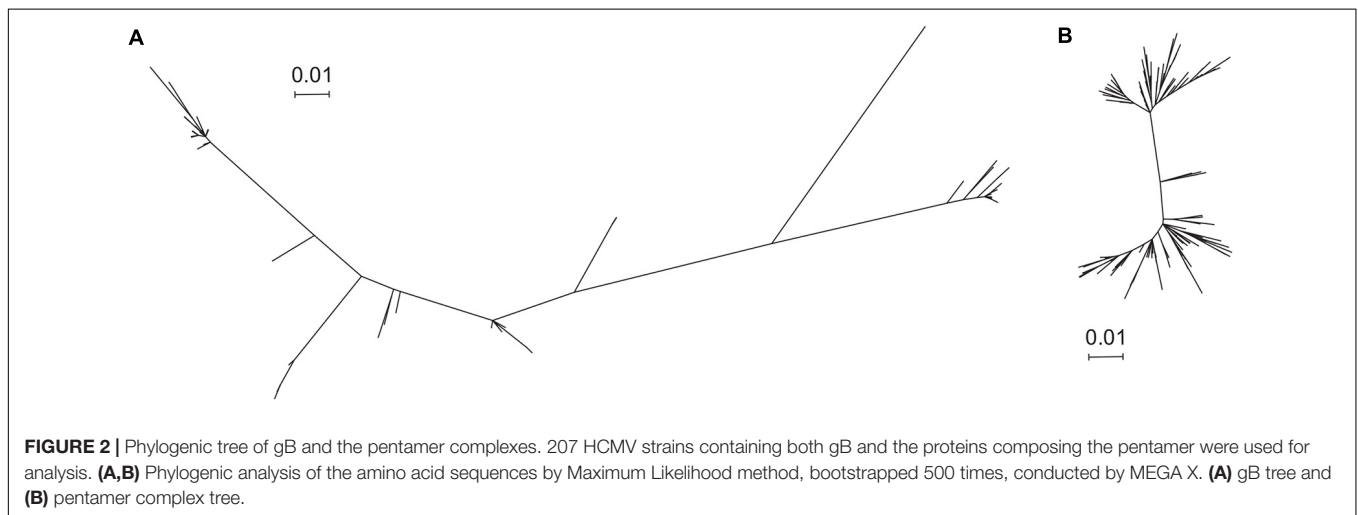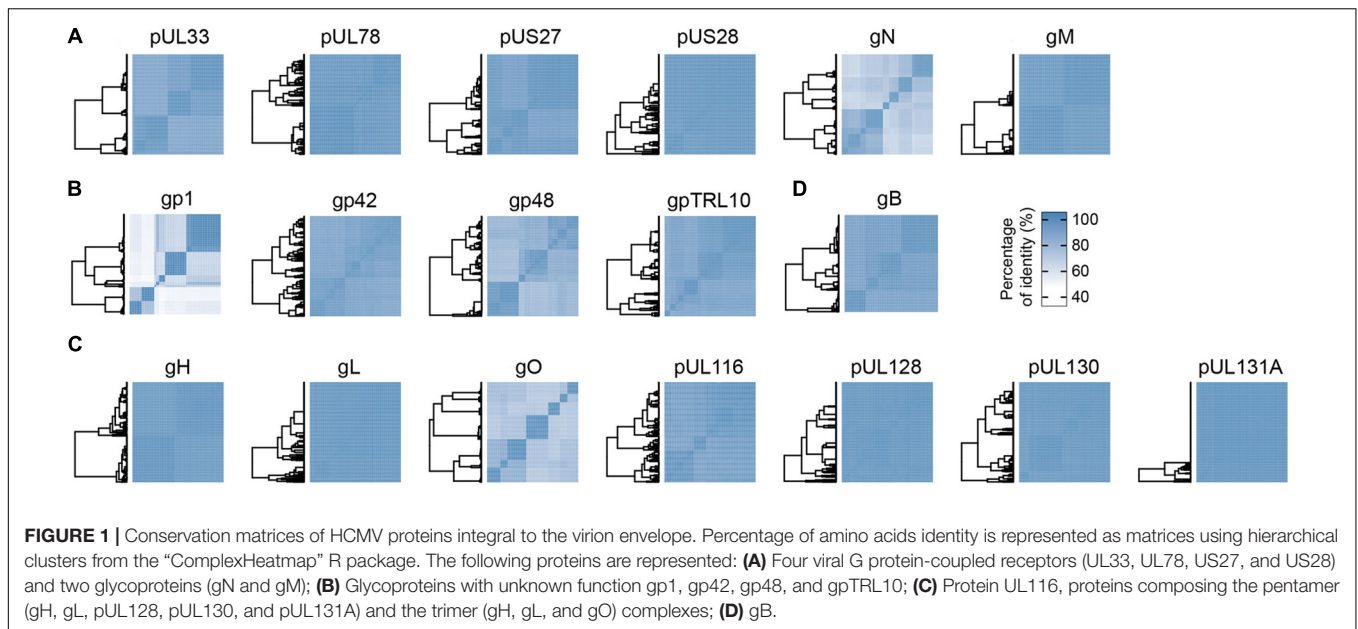## Sequence Diversity of HCMV Glycoproteins Involved in Viral Entry

Human cytomegalovirus uses two viral ligands for specific and high affinity interactions with the host cells. The pentamer complex is required for viral entry in epithelial, endothelial and myeloid cells (Ryckman et al., 2008a,b), and it represents the main target of neutralization against HCMV, eliciting extremely potent neutralizing antibodies (Macagno et al., 2010; Kabanova et al., 2014). Sequence analysis of the pentamer subunits reveals an extremely high level of conservation, with a mean identity of 98% (**Figure 1C**). This degree of conservation could be explained by the complex folding required to assemble the five subunits together (Chandramouli et al., 2017). Moreover, the multiple amino acid contacts occurring between the Nrp2 receptor and the gL, pUL128, pUL130, and pUL131A subunits also limit the possibility of antigenic drift (Martinez-Martin et al., 2018). Surprisingly, the trimeric complex composed of gH, gL, and gO possesses a low mean percentage of identity. The gO subunit is responsible for this low degree of conservation, since it is one of the less conserved envelope glycoproteins of HCMV, with a mean identity of 81% (**Figure 1C**). Interestingly, most of the sequence divergence is located in the N-terminus part of the protein (**Supplementary Figure S1b**), which is supposed to interact with PDGFRα, the cellular receptor of this complex (Stegmann et al., 2017, 2019). Of note, the pUL116 protein was

recently identified as a binding partner for gH, forming a dimeric complex of unknown function (Calo et al., 2016). The relatively high degree of conservation observed for pUL116 might indicate that the gH/pUL116 complex interacts with host or other viral proteins. The gB protein is encoded by the highly polymorphic gene UL55 (Paterson et al., 2002). However, our sequence analysis revealed a mean identity of 96% at the amino acid level, indicating that the protein is well conserved across strains (**Figure 1D**). The glycoprotein B is necessary for entry in all cell types, making it an appealing candidate for a HCMV vaccine.

Sequence Diversity of Vaccine Candidates

The trimer gHgLgO is necessary for cellular entry in all cell types and represents a potential target for vaccine design. Nonetheless, the complex glycan shield and the important sequence variation of gO are likely to represent a major problem for vaccine development. The gHgL heterodimer represents an interesting option for vaccine design, since both subunits are well conserved and characterized. However, immunization experiments in animal models generated suboptimal neutralization titers (Kabanova et al., 2014). Therefore, the gB homotrimer and the pentamer represent the best option for a prophylactic HCMV vaccine either alone or in combination. Some reports suggested that gB or pentamer sequence variation could be an issue for the generation of an efficient vaccine (Renzette et al., 2011). To gain further insights into gB and pentamer sequence variability, we used a total of 207 sequences per protein, obtained from the NCBI Entrez Protein database corresponding to different strains. This allowed us to generate a phylogenic tree based on the Maximum Likelihood method for both gB (**Figure 2A** and **Supplementary Figure S2**) and the pentamer (**Figure 2B** and **Supplementary Figure S3**). This analysis demonstrated that gB sequences are more diverse among strains (**Figure 2A** and **Supplementary Data S1**) compared to the pentamer (**Figure 2B** and **Supplementary Data S2**).

Indeed, analysis of sequence alignment demonstrated that the pentamer is well conserved, with only one main sequence variation localized on a short stretch of 30 amino acids (7 to 37) (**Figure 3A**) corresponding to the gH signal peptide (Chandramouli et al., 2017). In contrast, sequence variations on gB were localized on two "hot spots" (**Figure 3B**). These regions correspond to amino acids 27 to 70 and amino acids 451 to 483. The distal region (aa 451–483) includes the Furin cleavage site (Singh and Compton, 2000), and sequence variation of this site was initially used to classify different gB genotypes (Chou and Dennison, 1991). However, the relevance of the latter is not entirely clear concerning the infectivity and classification of HCMV strains (Stangherlin et al., 2017). The proximal region (aa 27–70) includes AD-2, which is one of the five antigenic domains (AD-1, -2, -3, -4 and -5) identified for gB. Interestingly, only antibodies directed against AD-2, AD-4, or AD-5 are able to neutralize the virus. Recently, a publication demonstrated that anti-AD-2 serum titers correlate with protection from viremia (Baraniak et al., 2018b), information that is highly relevant for vaccine design. Surprisingly, the crystal structure obtained

FIGURE 1 | Conservation matrices of HCMV proteins integral to the virion envelope. Percentage of amino acids identity is represented as matrices using hierarchical clusters from the "ComplexHeatmap" R package. The following proteins are represented: **(A)** Four viral G protein-coupled receptors (UL33, UL78, US27, and US28) and two glycoproteins (gN and gM); **(B)** Glycoproteins with unknown function gp1, gp42, gp48, and gpTRL10; **(C)** Protein UL116, proteins composing the pentamer (gH, gL, pUL128, pUL130, and pUL131A) and the trimer (gH, gL, and gO) complexes; **(D)** gB.

FIGURE 2 | Phylogenic tree of gB and the pentamer complexes. 207 HCMV strains containing both gB and the proteins composing the pentamer were used for analysis. **(A,B)** Phylogenic analysis of the amino acid sequences by Maximum Likelihood method, bootstrapped 500 times, conducted by MEGA X. **(A)** gB tree and **(B)** pentamer complex tree.

DISCUSSION

Human cytomegalovirus is the most common cause of viral-induced birth defects in industrialized countries. There is no treatment or vaccine clinically available today. The virion envelope protein gB is essential for viral entry and represents a major target of the humoral immune response following infection. gB vaccines were found to be safe and immunogenic in clinical phase I and II studies, eliciting an IgG response with a gB-binding magnitude comparable to that of a natural

from gB in its post-fusion conformation (Burke and Heldwein, 2015; Chandramouli et al., 2015) was not performed on the full extracellular domain and did not include the AD-2 sequence. The latter was removed, as it was believed to be unstructured and flexible in the post-fusion conformation adopted by gB.

infection (Baraniak et al., 2018a; Nelson et al., 2018b). However, despite the gB titers induced, only a minimal virus-neutralizing response was observed (Schleiss, 2018). The immunological reasons for these results are not fully understood yet, but several hypotheses have been proposed. The antigen employed comprised the intraluminal region of the gB molecule, a domain that generated non-neutralizing antibodies since it is not exposed onto the surface of the virion (Schleiss, 2018). The sequence of gB used in the vaccine is derived from a laboratory viral strain that possesses neither the tropism nor the virulence of clinical isolates (Nelson et al., 2018a). In addition, the sequence diversity observed at the genomic level was proposed as a potential issue for vaccine development (Renzette et al., 2017). Here, we performed sequence analyses for all the envelope proteins of HCMV and found that most of the proteins with a relevant function are well conserved, with the exception of gN and gO. While we investigated the relevance of antigenic variation for
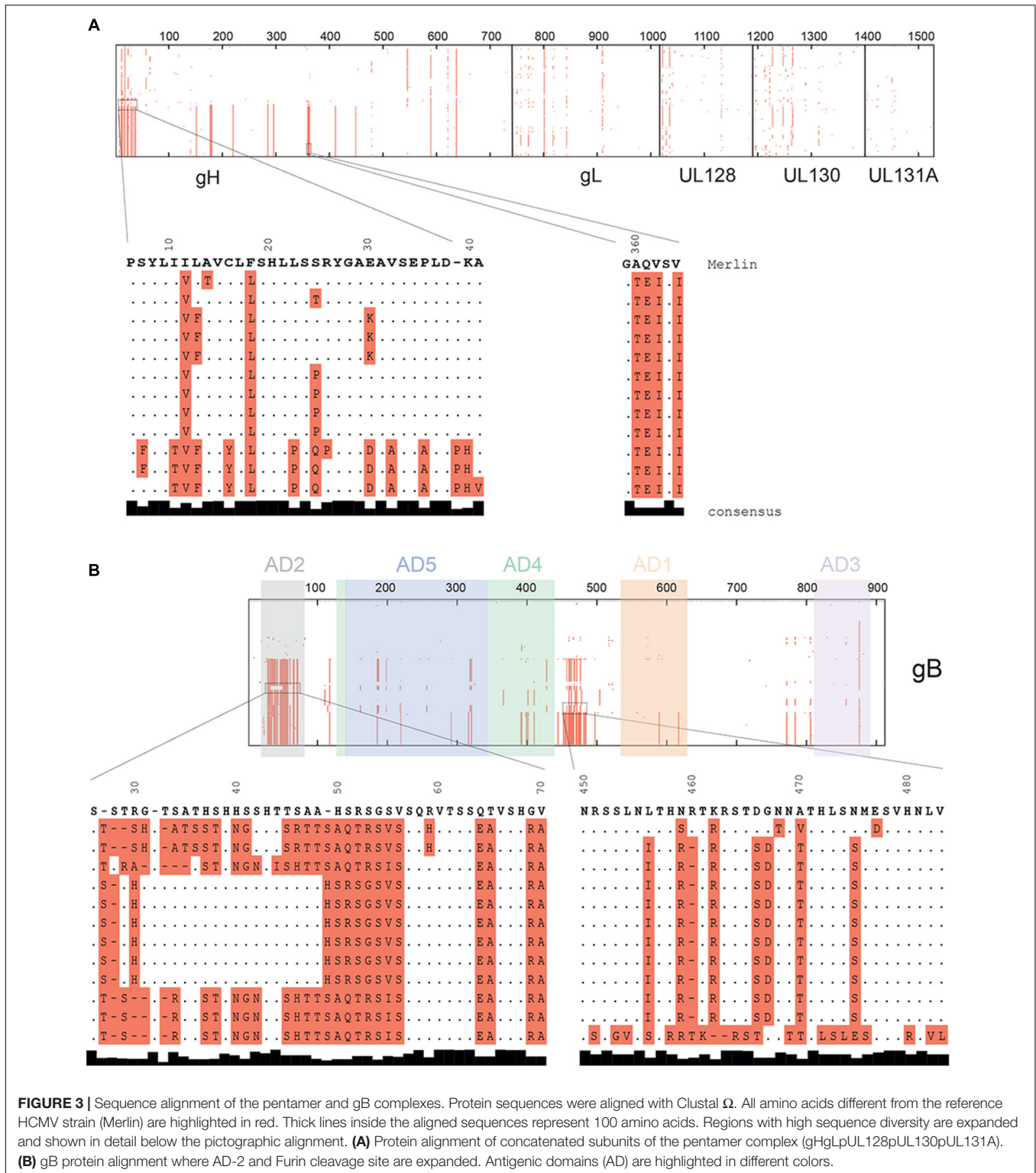
FIGURE 3 | Sequence alignment of the pentamer and gB complexes. Protein sequences were aligned with Clustal Ω. All amino acids different from the reference HCMV strain (Merlin) are highlighted in red. Thick lines inside the aligned sequences represent 100 amino acids. Regions with high sequence diversity are expanded and shown in detail below the pictographic alignment. **(A)** Protein alignment of concatenated subunits of the pentamer complex (gHgLpUL128pUL130pUL131A). **(B)** gB protein alignment where AD-2 and Furin cleavage site are expanded. Antigenic domains (AD) are highlighted in different colors.

vaccine design, we found that the pentamer is highly conserved across all HCMV strains with a deposited sequence. The gB antigen presents two main regions of variability across strains: the Furin cleavage site and part of AD-2. However, the Furin cleavage site is not supposed to be antigenic (Potzsch et al., 2011),

and only the conserved site of AD-2 is responsible for the generation of neutralizing antibodies (Baraniak et al., 2018b,c). Our analysis shows that both the pentamer and gB are well conserved across strains. We hypothesize that the modest efficacy observed in the recent clinical trials for an HCMV

vaccine (Baraniak et al., 2018a; Nelson et al., 2018b) does not result from HCMV genomic polymorphism. The absence of a strong neutralizing activity induced by the gB vaccine can be explained by the immunodominant response generated against AD-1 and AD-3. It has been shown that anti-AD1 antibodies do not correlate with protection (Baraniak et al., 2018b), and the humoral response against AD-1 upon natural infection generates a very modest neutralizing activity (Speckner et al., 1999). Moreover, AD-1 is also thought to mask AD-2, and part of the latter is the target of neutralizing antibodies (Baraniak et al., 2018b). Regarding AD-3, this domain is present in the lumen of the virus; therefore, as previously discussed, it is unlikely that antibodies generated against this domain will demonstrate a neutralizing activity (Schleiss, 2018). In summary, we speculate that engineering a gB vaccine lacking the antigenic domains AD-1 and AD-3 might generate a higher neutralizing titer in comparison to the vaccine candidate used in clinical trials.

DATA AVAILABILITY

All datasets analyzed for this study can be obtained by contacting the corresponding author.

AUTHOR CONTRIBUTIONS

MF designed and performed the bioinformatic analysis and wrote the manuscript. JM analyzed the data and wrote the manuscript. LP conceived the study, analyzed the data and wrote the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01005/full#supplementary-material

FIGURE S1 | Sequence alignment of gN and gO proteins. Amino acids different from the reference HCMV strain (Merlin) are highlighted in red. Thick lines inside the aligned sequences represent 100 amino acids. (**a**) Protein alignment of gN. (**b**) Protein alignment of gO.

FIGURE S2 | Labeled gB phylogenic tree. Phylogenic analysis of the amino acid sequences by Maximum Likelihood method for gB with name of the strains. Bootstrap values equal or higher than 70% are shown.

FIGURE S3 | Labeled pentamer phylogenic tree. Phylogenic analysis of the amino acid sequences by Maximum Likelihood method for the pentamer with name of the strains. Bootstrap values equal or higher than 70% are shown.

TABLE S1 | HCMV strains used for the study.

DATA S1 | gB alignment file in fasta format.

DATA S2 | Pentamer alignment file in fasta format.

REFERENCES

Alderete, J. P., Child, S. J., and Geballe, A. P. (2001). Abundant early expression of gpUL4 from a human cytomegalovirus mutant lacking a repressive upstream open reading frame. *J. Virol.* 75, 7188–7192. doi: 10.1128/jvi.75.15.7188-7192.2001

Anderholm, K. M., Bierle, C. J., and Schleiss, M. R. (2016). Cytomegalovirus vaccines: current status and future prospects. *Drugs* 76, 1625–1645. doi: 10.1007/s40265-016-0653-5

Baraniak, I., Kropff, B., Ambrose, L., McIntosh, M., McLean, G. R., Pichon, S., et al. (2018a). Protection from cytomegalovirus viremia following glycoprotein B vaccination is not dependent on neutralizing antibodies. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6273–6278. doi: 10.1073/pnas.1800224115

Baraniak, I., Kropff, B., McLean, G. R., Pichon, S., Piras-Douce, F., Milne, R. S. B., et al. (2018b). Epitope-specific humoral responses to human cytomegalovirus glycoprotein-B vaccine with MF59: anti-AD2 levels correlate with protection from Viremia. *J. Infect. Dis.* 217, 1907–1917. doi: 10.1093/infdis/jiy102

Baraniak, I. A., Reeves, M. B., and Griffiths, P. D. (2018c). Criteria to define interruption of transmission of human cytomegalovirus from organ donor to recipient. *Rev. Med. Virol.* 28:e1958. doi: 10.1002/rmv.1958

Burke, H. G., and Heldwein, E. E. (2015). Crystal structure of the human cytomegalovirus glycoprotein B. *PLoS Pathog* 11:e1005227. doi: 10.1371/journal.ppat.1005227

Calo, S., Cortese, M., Ciferri, C., Bruno, L., Gerrein, R., Benucci, B., et al. (2016). The human cytomegalovirus UL116 gene encodes an envelope glycoprotein forming a complex with gH independently from gL. *J. Virol.* 90, 4926–4938. doi: 10.1128/JVI.02517-15

Chandramouli, S., Ciferri, C., Nikitin, P. A., Calo, S., Gerrein, R., Balabanis, K., et al. (2015). Structure of HCMV glycoprotein B in the postfusion conformation bound to a neutralizing human antibody. *Nat. commun.* 6:8176. doi: 10.1038/ncomms9176

Chandramouli, S., Malito, E., Nguyen, T., Luisi, K., Donnarumma, D., Xing, Y., et al. (2017). Structural basis for potent antibody-mediated neutralization of human cytomegalovirus. *Sci. Immunol.* 2:eaan1457. doi: 10.1126/sciimmunol.aan1457

Chiuppesi, F., Kaltcheva, T., Meng, Z., Barry, P. A., Diamond, D. J., and Wussow, F. (2017). Identification of a continuous neutralizing epitope within UL128 of human cytomegalovirus. *J. Virol.* 91:e01857-16. doi: 10.1128/JVI.01857-16

Chou, S. W., and Dennison, K. M. (1991). Analysis of interstrain variation in cytomegalovirus glycoprotein B sequences encoding neutralization-related epitopes. *J. Infect. Dis.* 163, 1229–1234. doi: 10.1093/infdis/163.6.1229

Compton, T., Nowlin, D. M., and Cooper, N. R. (1993). Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate. *Virology* 193, 834–841. doi: 10.1006/viro.1993.1192

Congreve, M., Oswald, C., and Marshall, F. H. (2017). Applying structure-based drug design approaches to allosteric modulators of GPCRs. *Trends Pharmacol. Sci.* 38, 837–847. doi: 10.1016/j.tips.2017.05.010

de Munnik, S. M., Smit, M. J., Leurs, R., and Vischer, H. F. (2015). Modulation of cellular signaling by herpesvirus-encoded G protein-coupled receptors. *Front. Pharmacol.* 6:40. doi: 10.3389/fphar.2015.00040

Frascaroli, G., Lecher, C., Varani, S., Setz, C., van der Merwe, J., Brune, W., et al. (2018). Human macrophages escape inhibition of major histocompatibility complex-dependent antigen presentation by cytomegalovirus and drive proliferation and activation of memory CD4(+) and CD8(+) T Cells. *Front. Immunol.* 9:1129. doi: 10.3389/fimmu.2018.01129

Frascaroli, G., Varani, S., Moepps, B., Sinzger, C., Landini, M. P., and Mertens, T. (2006). Human cytomegalovirus subverts the functions of monocytes,

impairing chemokine-mediated migration and leukocyte recruitment. *J. Virol.* 80, 7578–7589. doi: 10.1128/jvi.02421-05

Galitska, G., Biolatti, M., De Andrea, M., Leone, A., Coscia, A., Bertolotti, L., et al. (2018). Biological relevance of Cytomegalovirus genetic variability in congenitally and postnatally infected children. *J. Clin. Virol.* 108, 132–140. doi: 10.1016/j.jcv.2018.09.019

Gao, J. L., and Murphy, P. M. (1994). Human cytomegalovirus open reading frame US28 encodes a functional beta chemokine receptor. *J. Biol. Chem.* 269, 28539–28542.

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847–2849. doi: 10.1093/bioinformatics/btw313

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275

Kabanova, A., Marcandalli, J., Zhou, T., Bianchi, S., Baxa, U., Tsybovsky, Y., et al. (2016). Platelet-derived growth factor-alpha receptor is the cellular receptor for human cytomegalovirus gHgLgO trimer. *Nat. Microbiol.* 1:16082. doi: 10.1038/nmicrobiol.2016.82

Kabanova, A., Perez, L., Lilleri, D., Marcandalli, J., Agatic, G., Becattini, S., et al. (2014). Antibody-driven design of a human cytomegalovirus gHgLpUL128L subunit vaccine that selectively elicits potent neutralizing antibodies. *Proc. Natl. Acad. Sci. U.S.A.* 111, 17965–17970. doi: 10.1073/pnas.1415310111

Kenneson, A., and Cannon, M. J. (2007). Review and meta-analysis of the epidemiology of congenital cytomegalovirus (CMV) infection. *Rev. Med. Virol.* 17, 253–276. doi: 10.1002/rmv.535

Kirby, T. (2016). Congenital cytomegalovirus-a neglected health problem. *Lancet Infect. Dis.* 16, 900–901. doi: 10.1016/s1473-3099(16)30226-2

Krishna, B. A., Miller, W. E., and O'Connor, C. M. (2018). US28: HCMV's swiss army knife. *Viruses* 10:445. doi: 10.3390/v10080445

Kropff, B., Koedel, Y., Britt, W., and Mach, M. (2010). Optimal replication of human cytomegalovirus correlates with endocytosis of glycoprotein gpUL132. *J. Virol.* 84, 7039–7052. doi: 10.1128/JVI.01644-09

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245. doi: 10.1093/nar/gkw290

Macagno, A., Bernasconi, N. L., Vanzetta, F., Dander, E., Sarasini, A., Revello, M. G., et al. (2010). Isolation of human monoclonal antibodies that potently neutralize human cytomegalovirus infection by targeting different epitopes on the gH/gL/UL128-131A complex. *J. Virol.* 84, 1005–1013. doi: 10.1128/JVI.01809-09

Malito, E., Chandramouli, S., and Carfì, A. (2018). From recognition to execution-the HCMV Pentamer from receptor binding to fusion triggering. *Curr. Opin. Virol.* 31, 43–51. doi: 10.1016/j.coviro.2018.05.004

Martinez-Martin, N., Marcandalli, J., Huang, C. S., Arthur, C. P., Perotti, M., Foglierini, M., et al. (2018). An unbiased screen for human cytomegalovirus identifies neuropilin-2 as a central viral receptor. *Cell* 174, 1158.e19–1171.e19. doi: 10.1016/j.cell.2018.06.028

Nelson, C. S., Herold, B. C., and Permar, S. R. (2018a). A new era in cytomegalovirus vaccinology: considerations for rational design of next-generation vaccines to prevent congenital cytomegalovirus infection. *NPJ Vacc.* 3:38. doi: 10.1038/s41541-018-0074-4

Nelson, C. S., Huffman, T., Jenks, J. A., Cisneros de la Rosa, E., Xie, G., Vandergrift, N., et al. (2018b). HCMV glycoprotein B subunit vaccine efficacy mediated by nonneutralizing antibody effector functions. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6267–6272. doi: 10.1073/pnas.1800177115

Nguyen, C. C., and Kamil, J. P. (2018). Pathogen at the gates: human cytomegalovirus entry and cell tropism. *Viruses* 10:E704. doi: 10.3390/v10120704

Paterson, D. A., Dyer, A. P., Milne, R. S., Sevilla-Reyes, E., and Gompels, U. A. (2002). A role for human cytomegalovirus glycoprotein O (gO) in cell fusion and a new hypervariable locus. *Virology* 293, 281–294. doi: 10.1006/viro.2001.1274

Pignatelli, S., Dal Monte, P., and Landini, M. P. (2001). gpUL73 (gN) genomic variants of human cytomegalovirus isolates are clustered into four distinct genotypes. *J. Gen. Virol.* 82, 2777–2784. doi: 10.1099/0022-1317-82-11-2777

Potzsch, S., Spindler, N., Wiegers, A. K., Fisch, T., Rucker, P., Sticht, H., et al. (2011). B cell repertoire analysis identifies new antigenic domains on glycoprotein B of human cytomegalovirus which are target of neutralizing antibodies. *PLoS Pathog* 7:e1002172. doi: 10.1371/journal.ppat.1002172

Rando, R. F., Srinivasan, A., Feingold, J., Gonczol, E., and Plotkin, S. (1990). Characterization of multiple molecular interactions between human cytomegalovirus (HCMV) and human immunodeficiency virus type 1 (HIV-1). *Virology* 176, 87–97. doi: 10.1016/0042-6822(90)90233-h

Renzette, N., Bhattacharjee, B., Jensen, J. D., Gibson, L., and Kowalik, T. F. (2011). Extensive genome-wide variability of human cytomegalovirus in congenitally infected infants. *PLoS Pathog* 7:e1001344. doi: 10.1371/journal.ppat.1001344

Renzette, N., Kowalik, T. F., and Jensen, J. D. (2016). On the relative roles of background selection and genetic hitchhiking in shaping human cytomegalovirus genetic diversity. *Mol. Ecol.* 25, 403–413. doi: 10.1111/mec.13331

Renzette, N., Pfeifer, S. P., Matuszewski, S., Kowalik, T. F., and Jensen, J. D. (2017). On the analysis of intrahost and interhost viral populations: human cytomegalovirus as a case study of pitfalls and expectations. *J. Virol.* 91:e01976-16. doi: 10.1128/JVI.01976-16

Ryckman, B. J., Chase, M. C., and Johnson, D. C. (2008a). HCMV gH/gL/UL128-131 interferes with virus entry into epithelial cells: Evidence for cell type-specific receptors. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14118–14123. doi: 10.1073/pnas.0804365105

Ryckman, B. J., Rainish, B. L., Chase, M. C., Borton, J. A., Nelson, J. A., Jarvis, M. A., et al. (2008b). Characterization of the human cytomegalovirus gH/gL/UL128-131 complex that mediates entry into epithelial and endothelial cells. *J. Virol.* 82, 60–70. doi: 10.1128/jvi.01910-07

Sayers, E. W., Agarwala, R., Bolton, E. E., Brister, J. R., Canese, K., Clark, K., et al. (2019). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 47, D23–D28. doi: 10.1093/nar/gky1069

Schleiss, M. R. (2018). Recombinant cytomegalovirus glycoprotein B vaccine: rethinking the immunological basis of protection. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6110–6112. doi: 10.1073/pnas.1806420115

Shikhagaie, M., Merce-Maldonado, E., Isern, E., Muntasell, A., Alba, M. M., Lopez-Botet, M., et al. (2012). The human cytomegalovirus-specific UL1 gene encodes a late-phase glycoprotein incorporated in the virion envelope. *J. Virol.* 86, 4091–4101. doi: 10.1128/JVI.06291-11

Shimamura, M., Mach, M., and Britt, W. J. (2006). Human cytomegalovirus infection elicits a glycoprotein M (gM)/gN-specific virus-neutralizing antibody response. *J. Virol.* 80, 4591–4600. doi: 10.1128/jvi.80.9.4591-4600.2006

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75

Singh, J., and Compton, T. (2000). Characterization of a panel of insertion mutants in human cytomegalovirus glycoprotein B. *J. Virol.* 74, 1383–1392. doi: 10.1128/jvi.74.3.1383-1392.2000

Speckner, A., Glykofrydes, D., Ohlin, M., and Mach, M. (1999). Antigenic domain 1 of human cytomegalovirus glycoprotein B induces a multitude of different antibodies which, when combined, results in incomplete virus neutralization. *J. Gen. Virol.* 80(Pt 8), 2183–2191. doi: 10.1099/0022-1317-80-8-2183

Stangherlin, L. M., de Paula, F. N., Icimoto, M. Y., Ruiz, L. G. P., Nogueira, M. L., Braz, A. S. K., et al. (2017). Positively selected sites at hcmv gb furin processing region and their effects in cleavage efficiency. *Front. Microbiol.* 8:934. doi: 10.3389/fmicb.2017.00934

Stegmann, C., Hochdorfer, D., Lieber, D., Subramanian, N., Stohr, D., Laib Sampaio, K., et al. (2017). A derivative of platelet-derived growth factor receptor alpha binds to the trimer of human cytomegalovirus and inhibits entry into fibroblasts and endothelial cells. *PLoS Pathog.* 13:e1006273. doi: 10.1371/journal.ppat.1006273

Stegmann, C., Rothemund, F., Sampaio, K. L., Adler, B., and Sinzger, C. (2019). The N-terminus of the human cytomegalovirus glycoprotein O is important for binding to the cellular receptor PDGFRalpha. *J. Virol.* [Epub ahead of print] doi: 10.1128/JVI.00138-19

Stern-Ginossar, N., Weisburd, B., Michalski, A., Le, V. T., Hein, M. Y., Huang, S. X., et al. (2012). Decoding human cytomegalovirus. *Science* 338, 1088–1093. doi: 10.1126/science.1227919

Tandon, R., and Mocarski, E. S. (2012). Viral and host control of cytomegalovirus maturation. *Trends Microbiol.* 20, 392–401. doi: 10.1016/j.tim.2012.04.008

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033

Weekes, M. P., Tomasec, P., Huttlin, E. L., Fielding, C. A., Nusinow, D., Stanton, R. J., et al. (2014). Quantitative temporal viromics: an approach to investigate host-pathogen interaction. *Cell* 157, 1460–1472. doi: 10.1016/j.cell.2014.04.028

Wussow, F., Chiuppesi, F., Meng, Z., Martinez, J., Nguyen, J., Barry, P. A., et al. (2018). Exploiting 2A peptides to elicit potent neutralizing antibodies by a multi-subunit herpesvirus glycoprotein complex. *J. Virol. Methods* 251, 30–37. doi: 10.1016/j.jviromet.2017.10.006

Zhuravskaya, T., Maciejewski, J. P., Netski, D. M., Bruening, E., Mackintosh, F. R., and St Jeor, S. (1997). Spread of human cytomegalovirus (HCMV) after infection of human hematopoietic progenitor cells: model of HCMV latency. *Blood* 90, 2482–2491.