

Cytomegalovirus immediate early gene UL37 encodes a novel MHC-like protein

Lucjan S. Wyrwicz¹✉ and Leszek Rychlewski²

¹Department of Gastroenterology, Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, Warszawa, Poland; ²BioInfoBank Institute, Poznań, Poland

Received: 15 December, 2007; revised: 12 February, 2008; accepted: 26 February, 2008
available on-line: 04 March, 2008

The cytomegalovirus (CMV) genome encodes four clusters of genes expressed immediately after infection — i.e.: UL36–38, UL122–123, TRS1–IRS1, and US3. The general function of these genes is associated with inhibition of cellular mechanisms of antiviral response. Although several biological processes have been mapped onto specific gene products, the knowledge of the molecular mechanism of their activity remains fragmentary. Here, we report the application of protein structure prediction methods in assigning the function to a glycosylated domain encoded by UL37 of CMV (gpUL37, UL37x3). The discerned similarity clearly points out that this domain represents a novel type of a major histocompatibility complex (MHC)-like protein, and consequently may play a central role in an additional mechanism of escape from antiviral response.

Keywords: *Betaherpesvirinae*, CMV, apoptosis, MHC I class, protein structure prediction, bioinformatics

INTRODUCTION

Human cytomegalovirus (HCMV) is an important agent of gastrointestinal infections and other diseases such as pneumonitis, hepatitis, and retinitis (Whitley, 1996), especially in immunocompromised patients (Steininger, 2007). HCMV is also one of the major viral agents associated with congenital disorders (Kenneson & Cannon, 2007). The pathogen, similarly to other species of *Herpesviridae*, has an ability to establish latent infection (Sinclair & Sissons, 2006) and exhibits various mechanisms preventing the apoptosis of infected cells (Andoniou & Degli-Esposti, 2006). The natural host cell range of human CMV is narrow and is mainly restricted to terminally differentiated epithelial, fibroblast and endothelial cells, as well as macrophages.

The human cytomegalovirus genome is a double-stranded DNA molecule of nearly 230 thousand base pairs (kbp), encoding more than 150 pro-

teins. During productive infection, HCMV genes are expressed in a three-step cascade, with phases designated as immediate-early (IE), early and late. The genes expressed immediately after infection (immediate early gene expression) are clustered in the genome in four distinct loci: UL36–38, UL122–123, TRS1–IRS1, and US3 (Colberg-Poley, 1996). The general function of these genes is associated with inhibition of cellular mechanisms of antiviral response and temporal regulation of gene expression. The IE genes are involved in the cell cycle block in a p53-dependent manner (UL122–123), regulation of antiviral response executed by PKR kinase (TRS1, IRS1), as well as in inhibition of apoptosis triggering (UL36, N-terminal segment of UL37) and deactivation of viral antigen presentation *via* modulation of expression of MHC class I proteins (US3) (Colberg-Poley, 1996). The immediate early proteins encoded by UL122 and UL123 (IE1 and IE2, respectively) play a central role in initiation and maintenance of

✉Corresponding author: Lucjan S. Wyrwicz, Department of Gastroenterology, Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, W. C. Roentgena 5, 02-781 Warszawa, Poland; tel: (48) 22 546 2933; fax: (48) 22 644 0209; e-mail: lucjan@bioinfo.pl

Abbreviations: IE, immediate early; CMV, Cytomegalovirus; SCOP, Structural Classification of Proteins; vMIA, viral mitochondria-localized inhibitor of apoptosis; PDB, Protein Data Bank; kbp, kilobase pairs; ORF, open reading frame.

HCMV gene expression in both lytic and latent infection (Stenberg & Stinski, 1985). The presence of analogous genomic loci has been confirmed for all sequenced related viruses of *Betaherpesvirinae*, although a notable variance in the gene content was reported (Knipe *et al.*, 2001).

The expression of genes from the UL36–UL38 region, followed by alternative splicing, results in the synthesis of several gene transcripts with a repertoire of corresponding protein products, whose individual functions have not been fully described yet (Colberg-Poley *et al.*, 2000). Of these, UL36 encodes a nonglycosylated cytoplasmic protein of anti-apoptotic activity, due to its role as an inhibitor of a caspase cascade (Andoniou & Degli-Esposti, 2006). The analysis of conservation of UL36 across a spectrum of CMV genotypes suggests that this protein is mutated in many important laboratory strains, including the most commonly used strain AD169 (Patterson & Shenk, 1999). A recent study by Terhune *et al.*, (2007) provided a strong evidence that yet another gene product of this gene cluster — pUL38 — is a cell-death inhibitor enabling efficient virus replication.

The dominant IE gene product involved in anti-apoptotic regulation is encoded by the UL37 gene (Goldmacher, 2002), spanning 2.7 kbp of the viral genome (Dolan *et al.*, 2004). Lee *et al.* (2000) in their study on mutation of a murine homologue (M37) demonstrated that this protein is a virulence factor, required for virulence *in vivo*, but not essential for replication *in vitro*. Other previous studies clearly pointed out that the protein acts as a negative regulator of mitochondrial activation of apoptosis (viral mitochondria-localized inhibitor of apoptosis; vMIA (McCormick *et al.*, 2003)). In turn, functional assays suggested that the anti-apoptotic activity of UL37 is achieved at a point downstream of caspase 8, but prior to cytochrome *c* release from mitochondria in a Bcl-2-like manner. This was noted despite a lack of traceable sequence similarity of UL37 to known cellular proteins with such activity (Andoniou & Degli-Esposti, 2006). In the mitochondrion the protein product seems to sequester Bax in the form of a vMIA–Bax complex (Arnoult *et al.*, 2004; Poncet *et al.*, 2004). Also by altering the mitochondrial bioenergetics vMIA causes a widely expressed cytopathic effect (Poncet *et al.*, 2006). A recent molecular modelling study on vMIA was augmented with mutational assays to reveal the potential mechanism of vMIA's molecular activity. Applying the techniques of fold recognition resulted in a potential model of the three-dimensional structure of vMIA as a viral homologue of Bcl-X(L) (Pauleau *et al.*, 2007).

The minimal fragment of UL37 exhibiting vMIA activity was mapped to the cleaved N-terminal fragment, encoded by the first exon (pUL37x1; (An-

doniou & Degli-Esposti, 2006)). Apparently, a major fraction of the UL37 open reading frame encodes a highly glycosylated protein (gpUL37; UL37x3) transported through the secretory pathway, *via* endoplasmic reticulum and the Golgi apparatus. The molecular function and impact on viral pathogenesis of this predominant fraction of UL37 open reading frame (gpUL37) is (as of now) undetermined (Andoniou & Degli-Esposti, 2006).

The *Herpesviridae* constitute a relatively old group of viruses, for which phylogenetic studies postulate an evolution parallel to that of the vertebrate host species (Montague & Hutchison, 2000). Thus, a significant fraction of the herpesvirus gene pool encompasses divergent homologues of host cellular proteins, adapted to host-specific and cell-type specific conditions (Holzerlandt *et al.*, 2002). The exact number of such genes is probably underestimated, due to analyses conducted on the level of primary protein structure (i.e. pairwise amino-acid sequence similarities). In our previous studies, we applied state-of-the-art methods for distant homology detection and combined them with protein structure prediction in order to study several divergent *Herpesviridae* proteins and outline the molecular background of their function. The approach was successfully applied in the case of the critical invasion factor glycoprotein L (gL) (Wyrwicz & Rychlewski, 2007b), the *Gammaherpesvirinae* transcription factor BcRF1 (Wyrwicz & Rychlewski, 2007c), the *Alphaherpesvirinae* transcriptional regulator ICP4 (Wyrwicz & Rychlewski, 2007a) and the neurovirulence protein UL45 (Wyrwicz *et al.*, 2008). Here, we present similar results of fold recognition and three dimensional modelling, as applied to *Betaherpesviridae* gpUL37 domain in order to postulate the molecular function of this immediate early glycoprotein.

MATERIALS AND METHODS

Assembly of UL37 family. The sequence of CMV UL37 (gi|9625722; strain AD169) was subjected to PSI-BLAST (Altschul *et al.*, 1997) searches against the NCBI NR (non-redundant) protein database (National Center for Biotechnology Information; NCBI; 01/11/2007). The search was conducted until profile convergence, with the cut-off for inclusion in the profile set equal to 0.001. Subsequently, the obtained set of UL37 homologues was clustered at 90% of sequence identity, using the CD-HIT tool (Li & Godzik, 2006). Finally, the resulting sequences were aligned with ClustalW (Thompson *et al.*, 1994) with minor manual corrections of the end result.

Annotation of predicted sequence features. The positions of potentially glycosylated asparagine residues were assigned using NetNGlyc server

(<http://www.cbs.dtu.dk/services/NetNGlyc/> (Gupta & Brunak, 2002)) and literature data (Hayajneh *et al.*, 2001a). Protein transmembrane segments were annotated with TMHMM tool (Sonnhammer *et al.*, 1998). The prediction of potential globular regions was performed with GlobPlot (Linding *et al.*, 2003).

Molecular modelling of the UL37 protein.

Sequences of the conserved globular domain of UL37 were subjected to the Structure Prediction Meta Server (<http://bioinfo.pl/meta> (Bujnicki *et al.*, 2001)). At this stage, secondary structure prediction was also obtained using PsiPred (Jones, 1999) and ProfSec (Rost & Sander, 1993), accessed *via* the Meta Server.

The resulting structural alignment was manually refined according to the quality assessment tool – Verify3D (Eisenberg *et al.*, 1997). Finally, homology models were obtained with Modeller version 6.2 (Sanchez & Sali, 2000) and screened with 3D-Jury – a consensus fold recognition method (Ginalski *et al.*, 2003).

RESULTS AND DISCUSSION

Various advanced distant homology mapping (Meta-Basic (Ginalski *et al.*, 2004), FFAS3 (Jaroszewski *et al.*, 2005)) and threading methods (3D-PSSM

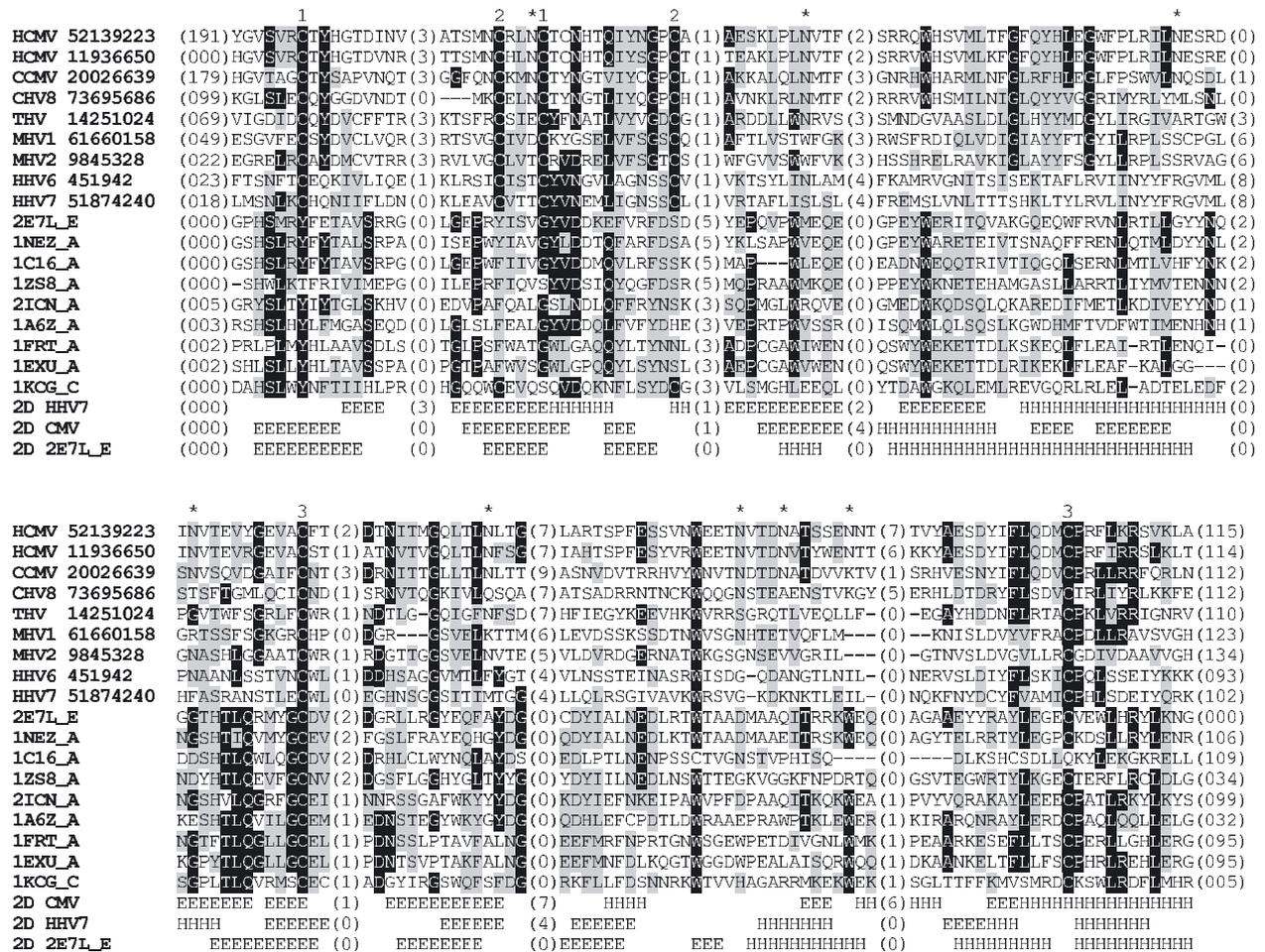


Figure 1. Alignment of UL37 glycoprotein and MHC protein domains.

Corresponding sequences were extracted from GenBank entries (coded with GenBank identifier gi; HCMV, *Human herpesvirus 5*; CCMV, *Chimpanzee cytomegalovirus*; CHV8, *Cercopithecine herpesvirus 8*; THV, *Tupaia herpesvirus*; MHV2, *Murid herpesvirus 2*; HHV6, *Human herpesvirus 6*; MHV1, *Murid herpesvirus 1*; HHV7, *Human herpesvirus 7*). Numbers in brackets refer to the length of removed fragments. Sequences of selected crystallographically solved MHC protein domains are included. The observed (PDB: 2E7L, chain E) and predicted (PsiPred, predictions for UL37 of CMV and HHV7) secondary structure elements are coded with letters (H, α -helix; E, β -strands). Location of potential N-glycosylation sites in CMV UL37, according to NetNGlyc tool, are marked with asterisk (*). The pairs of cystine residues creating disulphide bonds are denoted by corresponding numbers above the alignment. The PDB codes correspond to proteins of MHC-like fold: 2E7L, TCR receptor; 1NEZ, nonclassical MHC class I (TL protein); 1C16, nonclassical MHC class Ib (T22 protein); 1ZS8, nonclassical MHC class Ib (T10 protein); 2ICN, zinc α -2 glycoprotein; 1A6Z, hemochromatosis protein Hfe; 1FRT, 1EXU, Fc receptor; 1KCG, ULBP3.

(Kelley *et al.*, 2000), INUB (Fischer, 2003)) identified a major histocompatibility complex (MHC)-like protein fold as a potential structure of the investigated family (SCOP family d.19.1). As shown in Table 1, the prediction was confirmed with confident scores in the 3D-Jury method (Ginalski *et al.*, 2003). A summary of the top hits obtained for human pathogenic viruses of *Betaherpesvirinae* is shown in Table 2. Thus, we conclude that a consistent fold prediction was achieved for the conserved domain of UL37 proteins.

The predicted MHC fold consists of a β -sheet of eight strands and two long helices located at one side of the sheet in a $\beta\beta\beta\beta\alpha\beta\beta\beta\alpha$ topology (Garboczi *et al.*, 1996). The resulting domain has a hydrophobic ligand-binding cavity (located between the two helices). Most of the recognized MHC-fold proteins are involved in the binding of peptides in order to present either internal (Class I MHC) or external (Class II MHC) antigens in the process of acquired immune response. Here — the MHC domain is composed of two distinct fragments — referred to as α -1 and α -2 domains.

However, other ligands were also identified (Maenaka & Jones, 1999; Natarajan *et al.*, 1999). Although the binding module and the ligands differ between the observed functional groups of MHC proteins, a constant feature of this fold is the relatively high instability of the structure in the absence of the ligand (Garboczi *et al.*, 1996).

The structural alignment constructed on the basis of distant homology mappings, refined according to the predicted and observed secondary structures of gpUL37 and crystallographically solved proteins of MHC fold, respectively, is shown in Fig. 1. Although some discrepancy in the predictions of secondary structure is observed (compare Fig. 1), the overall topology of the proteins is conserved between the two protein families. A homology model of CMV gpUL37 built according to the alignment is shown in Fig. 2. Of note are the additional two pairs of conserved cysteine residues (marked with numbers above the alignment in Fig. 1), located in the model in spatial proximity. We postulate that these residues may stabilize the fold by both tightening the β -sheet and bridging one α -helix to the core part of the fold (Fig. 2).

The majority of proteins belonging to the MHC fold contain an additional all- β immunoglobulin-like domain, spacing the MHC fold from the cell membrane. In UL37 we also observe an additional segment located between the discussed domain and the transmembrane segment (Fig. 3). However, for this region (sometimes referred to as the basic domain (Hayajneh *et al.*, 2001b)), we were unable to notice homology to any other proteins of known fold.

The molecular mimicry of viral proteins to host proteins of defined functions is not unusual and several such events have been reported (Michel-

Table 1. Summary of fold recognition analysis for CMV UL37 (gi|52139223).

The top scoring hits (for each method) are shown in bold font. Hits below 3D-Jury cutoff of 50.0 (corresponding to less than 5% of prediction error) are shown in italics. Methods were coded according to their common abbreviations: 3DPS, 3D-PSSM; MBAS, MetaBasic; FFA3, FFAS3.

3DJury score	Method	# Selected Hit	Methods score	PDB identifier	Hit name	SCOP family	Organism
69.00	3DPS	3	0.171	1mhcA	MHC class I	d.19.1	R. norvegicus
68.71	3DPS	7	1.05	1zagA	Alpha-2-glycoprotein	d.19.1	H. sapiens
65.57	3DPS	1	0.142	1zagA	Alpha-2-glycoprotein	d.19.1	H. sapiens
62.57	INUB	3	10.93	1zagA	Alpha-2-glycoprotein	d.19.1	H. sapiens
60.43	3DPS	6	0.93	1duzA	MHC class I	d.19.1	H. sapiens
60.14	INUB	5	10.19	1kcgC	ULBP3 protein	d.19.1	H. sapiens
59.71	3DPS	4	0.252	1cd1A	CD1 protein	d.19.1	M. musculus
57.57	3DPS	8	1.46	1de4A	Hemochromatosis protein	d.19.1	H. sapiens
55.71	INUB	6	9.4	1jfmA	NK cell ligand RAE1 Beta	d.19.1	M. musculus
55.57	MBAS	2	21.39	1kcgC	ULBP3 protein	d.19.1	H. sapiens
54.29	INUB	1	14.05	1cd1A	CD1 protein	d.19.1	M. musculus
53.00	3DPS	9	1.55	1cd1A	CD1 protein	d.19.1	M. musculus
52.57	MBAS	8	18.44	1l8jA	Endothelial protein C receptor	d.19.1	H. sapiens
51.86	FFA3	1	-7.3	1jfmA	NK cell ligand RAE1 Beta	d.19.1	M. musculus
45.14	INUB	2	12.56	1a6zA	Hemochromatosis protein	d.19.1	H. sapiens
41.67	3DPS	5	0.757	3fruA	Neonatal Fc receptor	d.19.1	R. norvegicus
38.43	FFA3	2	-6.1	1l8jA	Endothelial protein C receptor	d.19.1	H. sapiens

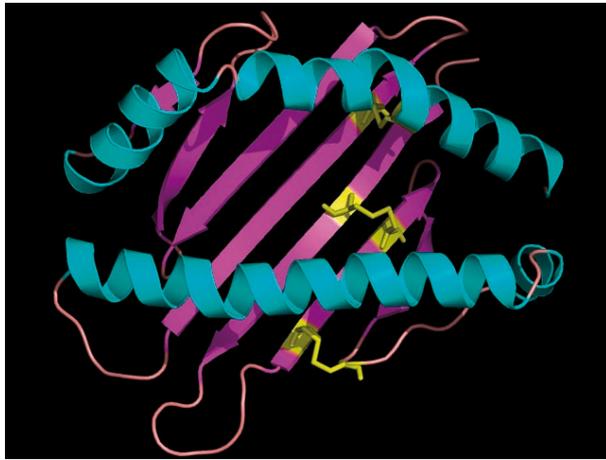


Figure 2. A model of CMV UL37 MHC-like domain. Secondary structure elements are coded with colors: cyan, α -helix; magenta, β -strands. Location of cysteine residues creating disulphide bonds is shown in yellow.

son, 2004). Also, the presence of a MHC class I homologue in the CMV genome, encoded by the UL18 gene, was previously reported (Beck & Barrell, 1988), further corroborating our present findings.

In the light of our analysis — the challenging question remains: “can gpUL37 bind any ligand?”. The modelling performed for such a low sequence similarity (below 10%) cannot be very helpful in determining the ligand-binding specificity (as this is a case of quite distant fold mapping).

Another open area is the reason for the wide preservation of MHC-like proteins in viral genomes of *Betaherpesvirinae*. The development of acquired immune system in early vertebrates, resulted in the creation of a complex MHC-fold-mediated antigen presentation and recognition system, based on a set of specific protein–protein interactions. All analyzed genomes of *Betaherpesvirinae* contain a homologue of UL37 and this subfamily diverged from other *Her-*

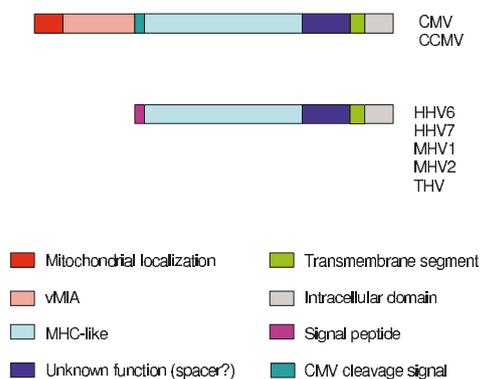


Figure 3. Overview of domain organization in UL37 orthologs (for virus abbreviations refer to Fig. 1).

pesviridae early (when compared to the evolution of the MHC protein family). Thus, we might speculate that this particular IE gene was acquired by an ancestor of *Betaherpesvirinae* in the moment of the divergence of MHC-fold proteins in the host genome.

The gene product of UL37 has been suggested to provide an important anti-apoptotic mechanism, one active early in the infection. This observation was consistently proven by several research groups in various experimental models and the reliability of this functional assignment is very high (Andoniou & Degli-Esposti, 2006). The UL37 homologues from the majority of *Betaherpesvirinae* genomes have a constant fragment of the open reading frame (ORF) encoded by CMV exon 3 (i.e., gpUL37), with the anti-apoptotic molecule vMIA (encoded by exon 1) restricted to primates’ CMV (HCMV, CCMV — compare Fig. 3). On this basis, we posit that vMIA was gained in the recent evolution of this *Cytomegalovirus* subfamily by inclusion of an additional exon in the ancestral IE gene. This information is also partially confirmed by the fact that both human and chimp CMV retain an ancestor signal peptide, located in the internal part of ORF, adjacent to the proposed MHC-like domain. This signal peptide is processed after the cleavage of vMIA (Mavinakere & Colberg-Poley, 2004). In such a situation, the functions of secreted gpUL37 and vMIA directed to the mitochondria can be synergic only in terms of a temporal association during IE gene expression.

The observed low sequence conservation of the glycosylated domain of UL37 (gpUL37) can be a result of viral adaptation to host response factors. Notably, this adaptation is not only driven by the host, but may also result from the presence of specific factors of infected cells (tissues) or viral pathogenesis. The glycosylated domain of UL37 from the human pathogens analyzed exhibits only 7–9% of sequence identity (in comparison of CMV *versus* closely related HHV6 and HHV7 sequences). This is consistent with the average sequence identity observed across the whole group of gpUL37 domain homologues. Additionally, none of the several potential N-glycosylation sites throughout the domain is conserved in all analyzed species (compare Fig. 1). This observation is consistent with a study showing relatively high sequence variance within primary isolates from clinical samples. A number of reported polymorphisms were located at potential N-glycosylation sites (Hayajneh *et al.*, 2001a), but the majority of these variants occurred at suboptimal N-glycosylation sites (according to the results of NetNGlyc method; not shown).

Since the binding of a ligand in the cavity of the MHC-like fold of gpUL37 remains disputable and we posit the existence of cysteine bonds bridging this unstable protein fold, we suggest that this

Table 2. Summary of results for fold recognition on the conserved domain of UL37 gene product, according to 3D-Jury prediction assessment method.

Template codes are Protein Data Bank (PDB) entry identifiers.

Virus	GenBank sequence identifier	3D-Jury score	Method	Template	SCOP
CMV	gi 52139223	69.00	3D-PSSM	1mhcA	d.19.1
HHV6A	gi 451942	44.25	MetaBasic	1l8jA	d.19.1
HHV6B	gi 9633087	43.50	3D-PSSM	1hocA	d.19.1
HHV7	gi 51874240	71.50	FFAS3	2e7IE	d.19.1

protein acts as a viral analog of the cellular proteins of the MHC complex (Kluczyk *et al.*, 2004). Their postulated function is to interfere with either class I or class II MHC proteins, rather than being an active member of antigen presentation mechanisms. Therefore, gpUL37 may define a novel robust viral mechanism of inactivation of the cell antiviral response. In this, it may cooperate with the other CMV proteins observed in *Betaherpesvirinae*, e.g.: US3 — a small protein ligand blocking the antigen presentation by binding to MHC class II proteins (Johnson & Hegde, 2002). Alternatively, gpUL37 can function similarly to a different MHC-like protein, encoded by the CMV gene UL18 (Beck & Barrell, 1988; Lopez-Botet *et al.*, 2001). The later gene product (pUL18) alters the activity of NK lymphocytes and dendritic cells (Wagner *et al.*, 2008).

In light of the above, further experimental assays (focused on mechanisms hypothesized here) are needed to elucidate the exact role of the glycosylated domain of UL37 homologues, during pathogenesis of *Betaherpesvirinae* infections.

Acknowledgements

This work was supported by Ministry of Science and Higher Education and the European Commission (COOP-CT-2004-508399) and MNiSW (2P05A00130, N401 050 32/1181). L.S.W. is supported by the Focus program from the Foundation for Polish Science.

REFERENCES

- Altschul S, Madden T, Schäffer A, Zhang J, Zhang Z, Miller W, Lipman D (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.
- Andoniou CE, Degli-Esposti MA (2006) Insights into the mechanisms of CMV-mediated interference with cellular apoptosis. *Immunol Cell Biol* **84**: 99–106.
- Arnoult D, Bartle LM, Skaletskaya A, Poncet D, Zamzami N, Park PU, Sharpe J, Youle RJ, Goldmacher VS (2004) Cytomegalovirus cell death suppressor vMIA blocks Bax- but not Bak-mediated apoptosis by binding and sequestering Bax at mitochondria. *Proc Natl Acad Sci USA* **101**: 7988–7993.
- Beck S, Barrell BG (1988) Human cytomegalovirus encodes a glycoprotein homologous to MHC class-I antigens. *Nature* **331**: 269–272.
- Bujnicki J, Elofsson A, Fischer D, Rychlewski L (2001) Structure prediction Meta server. *Bioinformatics* **17**: 750–751.
- Colberg-Poley AM (1996) Functional roles of immediate early proteins encoded by the human cytomegalovirus UL36–38, UL115–119, TRS1/IRS1 and US3 loci. *Intervirology* **39**: 350–360.
- Colberg-Poley AM, Patel MB, Erez DP, Slater JE (2000) Human cytomegalovirus UL37 immediate-early regulatory proteins traffic through the secretory apparatus and to mitochondria. *J Gen Virol* **81**: 1779–1789.
- Dolan A, Cunningham C, Hector RD, Hassan-Walker AF, Lee L, Addison C, Dargan DJ, McGeoch DJ, Gatherer D, Emery VC, Griffiths PD, Sinzger C, McSharry BP, Wilkinson GW, Davison AJ (2004) Genetic content of wild-type human cytomegalovirus. *J Gen Virol* **85**: 1301–1312.
- Eisenberg D, Luthy R, Bowie JU (1997) VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol* **277**: 396–404.
- Fischer D (2003) 3D-SHOTGUN: a novel, cooperative, fold-recognition meta-predictor. *Proteins* **51**: 434–441.
- Garboczi DN, Ghosh P, Utz U, Fan QR, Biddison WE, Wiley DC (1996) Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* **384**: 134–141.
- Ginalski K, Elofsson A, Fischer D, Rychlewski L (2003) 3D-Jury: a simple approach to improve protein structure predictions. *Bioinformatics* **19**: 1015–1018.
- Ginalski K, von Grotthuss M, Grishin N, Rychlewski L (2004) Detecting distant homology with Meta-BASIC. *Nucleic Acids Res* **32**: W576–W581.
- Goldmacher VS (2002) vMIA, a viral inhibitor of apoptosis targeting mitochondria. *Biochimie* **84**: 177–185.
- Gupta R, Brunak S (2002) Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput* 310–322.
- Hayajneh WA, Colberg-Poley AM, Skaletskaya A, Bartle LM, Lesperance MM, Contopoulos-Ioannidis DG, Kedersha NL, Goldmacher VS (2001a) The sequence and antiapoptotic functional domains of the human cytomegalovirus UL37 exon 1 immediate early protein are conserved in multiple primary strains. *Virology* **279**: 233–240.
- Hayajneh WA, Contopoulos-Ioannidis DG, Lesperance MM, Venegas AM, Colberg-Poley AM (2001b) The carboxyl terminus of the human cytomegalovirus UL37 immediate-early glycoprotein is conserved in primary strains and is important for transactivation. *J Gen Virol* **82**: 1569–1579.
- Holzerlandt R, Orengo C, Kellam P, Alba MM (2002) Identification of new herpesvirus gene homologs in the human genome. *Genome Res* **12**: 1739–1748.

- Jaroszewski L, Rychlewski L, Li Z, Li W, Godzik A (2005) FFAS03: a server for profile-profile sequence alignments. *Nucleic Acids Res* **33**: W284–W288.
- Johnson DC, Hegde NR (2002) Inhibition of the MHC class II antigen presentation pathway by human cytomegalovirus. *Curr Top Microbiol Immunol* **269**: 101–115.
- Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**: 195–202.
- Kelley L, MacCallum R, Sternberg M (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J Mol Biol* **299**: 499–520.
- Kenneson A, Cannon MJ (2007) Review and meta-analysis of the epidemiology of congenital cytomegalovirus (CMV) infection. *Rev Med Virol* **17**: 253–276.
- Kluczyk A, Cebrat M, Zbozien-Pacamaj R, Lisowski M, Stefanowicz P, Wiczorek Z, Siemion IZ (2004) On the peptide-antipeptide interactions in interleukin-1 receptor system. *Acta Biochim Polon* **51**: 57–66.
- Knipe D, Howley P, Griffin D, Lamb R, Martin M, Roizman B, Straus S, eds. (2001) *Fields Virology*. Lippincott Williams & Wilkins, Philadelphia.
- Lee M, Xiao J, Haghjoo E, Zhan X, Abenes G, Tuong T, Dunn W, Liu F (2000) Murine cytomegalovirus containing a mutation at open reading frame M37 is severely attenuated in growth and virulence *in vivo*. *J Virol* **74**: 11099–11107.
- Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**: 1658–1659.
- Linding R, Russell R, Neduva V, Gibson T (2003) GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res* **31**: 3701–3708.
- Lopez-Botet M, Llano M, Ortega M (2001) Human cytomegalovirus and natural killer-mediated surveillance of HLA class I expression: a paradigm of host-pathogen adaptation. *Immunol Rev* **181**: 193–202.
- Maenaka K, Jones EY (1999) MHC superfamily structure and the immune system. *Curr Opin Struct Biol* **9**: 745–753.
- Mavinakere MS, Colberg-Poley AM (2004) Internal cleavage of the human cytomegalovirus UL37 immediate-early glycoprotein and divergent trafficking of its proteolytic fragments. *J Gen Virol* **85**: 1989–1994.
- McCormick AL, Smith VL, Chow D, Mocarski ES (2003) Disruption of mitochondrial networks by the human cytomegalovirus UL37 gene product viral mitochondrion-localized inhibitor of apoptosis. *J Virol* **77**: 631–641.
- Michelson S (2004) Consequences of human cytomegalovirus mimicry. *Hum Immunol* **65**: 465–475.
- Montague MG, Hutchison CA 3rd (2000) Gene content phylogeny of herpesviruses. *Proc Natl Acad Sci USA* **97**: 5334–5339.
- Natarajan K, Li H, Mariuzza RA, Margulies DH (1999) MHC class I molecules, structure and function. *Rev Immunogenet* **1**: 32–46.
- Patterson CE, Shenk T (1999) Human cytomegalovirus UL36 protein is dispensable for viral replication in cultured cells. *J Virol* **73**: 7126–7131.
- Pauleau AL, Larochette N, Giordanetto F, Scholz SR, Poncet D, Zamzami N, Goldmacher VS, Kroemer G (2007) Structure-function analysis of the interaction between Bax and the cytomegalovirus-encoded protein vMIA. *Oncogene* **26**: 7067–7080.
- Poncet D, Larochette N, Pauleau AL, Boya P, Jalil AA, Cartron PF, Vallette F, Schnebelen C, Bartle LM, Skaletskaya A, Boutolleau D, Martinou JC, Goldmacher VS, Kroemer G, Zamzami N (2004) An anti-apoptotic viral protein that recruits Bax to mitochondria. *J Biol Chem* **279**: 22605–22614.
- Poncet D, Pauleau AL, Szabadkai G, Voza A, Scholz SR, Le Bras M, Briere JJ, Jalil A, Le Moigne R, Brenner C, Hahn G, Wittig I, Schagger H, Lemaire C, Bianchi K, Souquere S, Pierron G, Rustin P, Goldmacher VS, Rizzuto R, Palmieri F, Kroemer G (2006) Cytopathic effects of the cytomegalovirus-encoded apoptosis inhibitory protein vMIA. *J Cell Biol* **174**: 985–996.
- Rost B, Sander C (1993) Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proc Natl Acad Sci USA* **90**: 7558–7562.
- Sanchez R, Sali A (2000) Comparative protein structure modeling. Introduction and practical examples with modeller. *Methods Mol Biol* **143**: 97–129.
- Sinclair J, Sissons P (2006) Latency and reactivation of human cytomegalovirus. *J Gen Virol* **87**: 1763–1779.
- Sonnhammer EL, von Heijne G, Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**: 175–182.
- Steininger C (2007) Clinical relevance of cytomegalovirus infection in patients with disorders of the immune system. *Clin Microbiol Infect* **13**: 953–963.
- Stenberg RM, Stinski MF (1985) Autoregulation of the human cytomegalovirus major immediate-early gene. *J Virol* **56**: 676–682.
- Terhune S, Torigoi E, Moorman N, Silva M, Qian Z, Shenk T, Yu D (2007) Human cytomegalovirus UL38 protein blocks apoptosis. *J Virol* **81**: 3109–3123.
- Thompson J, Higgins D, Gibson T (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.
- Wagner CS, Walther-Jallow L, Buentke E, Ljunggren HG, Achour A, Chambers BJ (2008) Human cytomegalovirus-derived protein UL18 alters the phenotype and function of monocyte-derived dendritic cells. *J Leukoc Biol* **83**: 56–63.
- Whitley RJ (1996). *Herpesviruses In: Medical Microbiology*. Galveston, The University of Texas Medical Branch at Galveston.
- Wyrwicz LS, Rychlewski L (2007a) Fold recognition insights into function of herpes ICP4 protein. *Acta Biochim Polon* **54**: 551–559.
- Wyrwicz LS, Rychlewski L (2007b) Herpes glycoprotein gL is distantly related to chemokine receptor ligands. *Antiviral Res* **75**: 83–86.
- Wyrwicz LS, Rychlewski L (2007c) Identification of Herpes TATT-binding protein. *Antiviral Res* **75**: 167–172.
- Wyrwicz LS, Ginalska K, Rychlewski L (2008) HSV-1 UL45 encodes a carbohydrate binding C-type lectin protein. *Cell Cycle* **7**: 269–271.