The Structure of the Cytomegalovirus-Encoded m04 Glycoprotein, a Prototypical Member of the m02 Family of Immunoevasins*

Received for publication, May 27, 2014, and in revised form, June 24, 2014. Published, JBC Papers in Press, June 30, 2014. DOI 10.1074/jbc.M114.584128

Richard Berry‡1,2, Julian P. Vivian†, Felix A. Deuss‡, Gautham R. Balaji†, Philippa M. Saunders§, Jie Lin§, Dene R. Little‡, Andrew G. Brooks§, and Jamie Rossjohn*†1,3

From the ‡Department of Biochemistry and Molecular Biology, School of Biomedical Sciences and the §Australian Research Council Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton, Victoria 3800, Australia, the †Department of Microbiology & Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Parkville, Victoria 3010, Australia, and the §Institute of Infection and Immunity, Cardiff University, School of Medicine, Heath Park, Cardiff CF14 4XN, United Kingdom

Background: Cytomegalovirus-encoded m02 family members are involved in immune evasion.

Results: m04 binds murine MHC-I via an Ig-V-like ectodomain.

Conclusion: m04 is the prototype of the m02 family, whose members mimic proteins derived from the host immune system.

Significance: The results provide insight into the structural scaffold that defines a large family of immunoevasins.

The ability of CMVs to evade the immune system of the host is dependent on the expression of a wide array of glycoproteins, many of which interfere with natural killer cell function. In murine CMV, two large protein families mediate this immune-evasive function. Although it is established that the m145 family members mimic the structure of MHC-I molecules, the structure of the m02 family remains unknown. The most extensively studied m02 family member is m04, a glycoprotein that escorts newly assembled MHC-I molecules to the cell surface, presumably to avoid “missing self” recognition. Here we report the crystal structure of the m04 ectodomain, thereby providing insight into this large immunoavasins family. m04 adopted a β-sandwich immunoglobulin variable (Ig-V)-like fold, despite sharing very little sequence identity with the Ig-V superfamily. In addition to the Ig-V core, m04 possesses several unique structural features that included an unusual β-strand topology, a number of extended loops and a prominent α-helix. The m04 interior was packed by a myriad of hydrophobic residues that form distinct clusters around two conserved tryptophan residues. This hydrophobic core was well conserved throughout the m02 family, thereby indicating that murine CMV encodes a number of Ig-V-like molecules. We show that m04 binds a range of MHC-I molecules with low affinity in a peptide-independent manner. Accordingly, the structure of m04, which represents the first example of an murine CMV encoded Ig-V fold, provides a basis for understanding the structure and function of this enigmatic and large family of immunoevasins.

Many viruses avoid detection by cytotoxic T cells of the adaptive immune system through strategies that interfere with the surface expression of peptide-MHC-I molecules. These include mechanisms to interfere with the flow and trafficking of MHC-I through the secretory pathway, that interrupt the supply of peptides to MHC-I in the ER, or that modulate the rate of MHC-I endocytosis (1, 2). However, the down-regulation of surface MHC-I renders infected cells susceptible to natural killer (NK) cell-mediated lysis by a process termed “missing self” recognition (2). Accordingly, many viruses also encode molecules that seek to manipulate the function of a number of large, typically inhibitory, NK cell receptor families that include the killer immunoglobulin receptors in humans, the Ly49 receptors in mice, and the NKG2 family in both species (3).

The interplay between viral defense and host adaptation is best understood in the context of CMV infection. CMVs are a group of species-specific pathogens that are known for their ability to subvert the immune system of their host and establish latent and lifelong infections (4). They are highly prevalent, infecting over 50% of the human population worldwide with infections rates ranging from 45 to 100% in human adult populations and, although asymptomatic in healthy individuals, can cause considerably morbidity and mortality if the immune system is compromised (5). The success of CMV in part results from the large size of its genome (~230 kb), which affords the virus the opportunity to dedicate considerable genetic material to encoding proteins whose function confers even a slender evolutionary advantage (6).

Murine CMV (MCMV), which represents the model system for human CMV infection, encodes two large gene clusters whose protein products have been implicated in the subversion of NK cell-mediated recognition (7). Of these, the m145 family, which resides at the extreme right of the MCMV genome and includes m17 and m145–m158, has been extensively studied.

* This work is supported by the Australian Research Council and the National Health and Medical Research Council of Australia, a National Health and Medical Research Council Peter Doherty Fellowship (to R. B.), an Australian Research Council Discovery Early Career Researcher Award (to J. R. V.), and a National Health and Medical Research Council Australia Fellowship (to J. R. V.).

† These authors contributed equally to this work.

‡ The atomic coordinates and structure factors (code 4PN6) have been deposited in the Protein Data Bank (http://wwpdb.org/).

‡1,2 To whom correspondence may be addressed. Tel.: +61-3-9902-9239; E-mail: Richard.berry@monash.edu.

‡ To whom correspondence may be addressed. Tel.: +61-3-9902-9236; Fax: 99054699; E-mail: Jamie.rossjohn@monash.edu.

4 The abbreviations used are: NK, natural killer; MCMV, murine CMV; β2m, β2-microglobulin; MALLS, multicycle laser light scattering.
The Crystal Structure of m04

So far, the structures of four members of the m145 family (m144, m152, m153, and m157) have been determined (8–11). Remarkably, despite employing diverse mechanisms to mediate their function, each of these molecules possesses an MHC-I-like fold, although they differ in their requirement for peptide and β2-microglobulin (β2m).

At the extreme left of the MCMV genome lie the less understood m02 family members that include m02–m16 (7). Although the function of most of the m02 family members remains unknown, as a group they have been implicated in the subversion of NK cell-mediated immunity. This function was originally attributed based on experiments in which viruses harboring a deletion of entire m02 family grew like their wild-type parental strain in cultured cells or in C57BL/6 mice depleted of NK cells but produced less infectious viruses in mice possessing a robust NK cell response (12).

The members of the m02 family can be clustered into three subgroups, m02–m06, m07–m10, and m11–m16 based on the presence of conserved cysteine residues (7). The family members also share a number of conserved motifs that include a CXXXXC motif (that is more stringently defined as CXLXXC(L/P)/W/R/o in m02–m06) and a NAXXX(V/E/H)/Wo motif found in m03–m06 (where X indicates any amino acid, and o indicates a hydrophobic amino acid) (7). Although the m02 family members are related in the amino acid sequence, they are not appreciably similar to any other MCMV-encoded protein or to any other gene in the GenBank™ database (12).

To date, two members of the m02 family have been implicated in MHC-I targeting. Both m04 and m06 bind to newly assembled peptide-MHC-I complexes in the ER. However, whereas m06 redirects MHC-I to the lysosome for degradation (13), m04-MHC-I complexes traverse to the cell surface (14). Precisely why MCMV encodes a protein that escorts MHC-I to the cell surface is unclear, but it may be that maintaining a low level of MHC-I on the cell surface is beneficial to avoid NK cell-mediated missing self-recognition. More recently it has become apparent that m04 is the target of a novel viral detection strategy. In particular, there is now evidence that a number of activating NK cell receptors including Ly49P, Ly49L, and Ly49W can recognize MCMV-infected cells of certain H2 haplotypes (including H2d, H2k, H2, and H2) in an m04-dependent manner (15, 16). However, the precise molecular details underpinning such a recognition event remain a mystery. To shed light on m04 function, here we determined the crystal structure of m04. The structure revealed that m04 adopted an Ig-V-like scaffold that provides a basis for understanding the structure and function of the m02 family.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—DNA encoding the full-length nucleotide sequence for m04 from the MCMV isolates Smith, G4, and W8211 was a kind gift from Anthony Scalzo (17). Fragments encoding the predicted m04 extracellular domains (residues 24–223 for m04Smith and m04G4 and residues 24–220 for m04W8211) were amplified by PCR and ligated into the Agel and Xhol sites of the pHLSec vector (18). The reverse primers used for PCR were designed to incorporate a thrombin site (LVPRGS) and His8 tag at the C terminus of the expressed protein. m04 protein was expressed using transient transfection of HEK 293S cells as described previously (18). Cell culture media containing secreted protein were concentrated and buffer-exchanged into 10 mM Tris, pH 8, containing 0.5 M NaCl using tangential flow filtration prior to purification using nickel affinity and size exclusion chromatography using Superdex S75 16/60 columns (GE Healthcare) in 10 mM Tris, pH 8, containing 150 mM NaCl. MHC-I molecules were expressed as inclusion bodies in BL21 DE3 Escherichia coli cells and were refolded and purified essentially as described (19). The following peptide MHC-I complexes were employed; H2-D4 (RGPGRAFVTI), H2-L4 (YPHYFMPPTNL), H2-D4 (RL8: RRLRGLTLL and RL9: RRLGRTLLL), and HLA-B*5701 (KAFSPEVIPMF).

Surface Plasmon Resonance—SPR experiments were conducted at 20 °C on a Biacore 3000 instrument using 10 mM Tris, pH 8.0, supplemented with 150 mM NaCl and 0.005% P20 surfactant. Approximately 1300 response units of biotinylated MHC-I molecules were coupled to streptavidin-coated chips (GE Healthcare) according to the manufacturer’s instructions and the remaining free streptavidin sites were blocked with d-biotin. Various concentrations of m04 (3.1–200 μM) were injected over the captured MHC-I at a flow rate of 20 μl min⁻¹. The final response was calculated by subtracting the response of an “empty” flow cell (containing biotin-blocked streptavidin). The equilibrium data were analyzed using GraphPad Prism. The data are representative of a single experiment performed in duplicate.

Crystallization and Data Collection—For crystallization experiments, the extracellular domain of m04G4 was concentrated to 15.5 mg/ml. Crystals were obtained using the hanging drop vapor diffusion method from a solution containing 23% PEG 3350 and 0.2 mM sodium malonate, pH 4, at 4 °C. Ytterbium derivative crystals were obtained by soaking crystals overnight in the presence of 20 mM ytterbium-(DPA)3 (Jena Bioscience). Prior to data collection, crystals were equilibrated in a crystallization solution supplemented with 35% PEG 3350 for cryoprotection. Crystals were flash cooled using liquid nitrogen, and x-ray diffraction data were recorded on a Quantum-315 CCD detector at the MX2 beamline of the Australian Synchrotron. Data sets to 3.0 Å (native) and 3.6 Å (ytterbium derivative) were integrated using MOSFLM and scaled using SCALA from the CCP4 program suite. Details of the data processing statistics are given in Table 1.

Structure Determination and Refinement—The structure of the extracellular domain of m04G4 was determined by the single wavelength anomalous dispersion method. Diffraction data containing anomalous signal from bound ytterbium atoms was collected at a single high energy remote wavelength (1.148 Å). Experimental phasing was performed using Phenix Autosol (20). A preliminary m04 model was built into the solvent-flattened electron density map de novo and then used as a molecular replacement model for the higher resolution 3 Å native data set using Phaser (21). Obtaining the correct polypeptide register was greatly facilitated by the presence of clear unbiased electron density for the three N-linked glycosylation sites at Asn54, Asn115, and Asn130 early during the refinement process. The model was refined using Buster to an Rfree/Rref of 22.1%/
To further our understanding of the function of m04, we expressed the m04 ectodomain from a molecular mass moment of 23,980 Da but agrees well with the mass estimated from SDS-PAGE analysis. This discrepancy is attributable to the presence of multiple N-linked glycan moieties that are visible at three of the five predicted NX(S/T) sites (Asn43, Asn115, and Asn130). Accordingly, m04 is monomeric in solution, although we cannot exclude the possibility of dimer formation at higher protein concentrations or on the cell surface.

**Overview of the m04 Structure**—The core of the m04 molecule comprises an 11-stranded antiparallel beta sandwich (Fig. 1B) with overall dimensions of \(50 \times 35 \times 30 \text{ Å}\). The strand assignment and their arrangement into sheets are shown in Fig. 1C. The two \(\beta\)-sheets that make up the sandwich are curved in appearance and comprise four and seven \(\beta\)-strands (sheet 1: BCGH, sheet 2: AA‘IDEE’F). Because strands A and E are interrupted by short bulged regions, we have denoted the two parts as A/E and A‘/E‘. Because of this discontinuity, the width of sheet 2 is limited to five \(\beta\)-strands (Fig. 1C). The overall fold of the \(\beta\)-sandwich is similar to that of the immunoglobulin superfamily; strands A, B, C, D, E, F, G, H, and L correspond to strands G, A, B, C, C’, C”, D, E, and F in an immunoglobulin variable domain (IgV). Moreover, the classical hallmark of an Ig domain, a single disulfide bond that connects the B and F strands, is present in an analogous location in m04 and is derived from Cys46 and Cys131 in the C-D loop and I strands, respectively. In addition to the Ig core, m04 contains a number of extended loops and three short 3_10 helices located between the D-E, F-G, and H-I strands. Moreover, the m04 C terminus contains a bent \(\alpha\)-helix of 12 amino acids in length that sits in a concave surface formed by the A’ and I strands and the C-D loop (Fig. 1B). The presence of this helix (termed \(\alpha\)4) gives the part of the molecule distal to the Ig domain a “mini platform”-like appearance, in which the C-D loop lies in-between and parallel to the E-F loop and \(\alpha\)4 helix and is supported by the AA‘IDEE’F sheet (Fig. 1D).

Although the core of m04 is rich in hydrophobic residues (discussed below), the surface of the molecule is primarily basic in nature, reflecting its relatively high theoretical isoelectric point (9.03) (Fig. 2A). The three N-linked glycosylation sites lay in a linear arrangement along a basic strip derived from arginine residues at positions 20 and 88 and lysines at positions 18, 106, and 108. One of the glycosylation sites (at Asn17) is present in a loop that is in a homologous location to complementarity...
**FIGURE 1. Overview of the m04 structure.** A, the refractive index of m04\(^{G4}\) (thin line) measured during size exclusion chromatography is shown along with the molecular mass (thick line) calculated from simultaneous MALLS analysis. B, cartoon representation of the m04 structure. Secondary structure elements are colored as follows: \(\beta\)-sheets in blue, helices in red, and \(3_{10}\) helices in orange. Asparagine residues to which N-linked glycans (blue sticks) are attached are displayed as black sticks. The electron density (2\(F_o\) − \(F_c\)) map surrounding the N-linked glycans is displayed as blue mesh and is contoured at 1\(\sigma\). Cysteine residues that participate in disulfide bonding are represented by green sticks. C, schematic representation of the m04 secondary structure topology. \(\beta\)-Strands are represented by arrows, and helices are represented by rectangles. D, view of the m04 mini platform with the \(\alpha_4\) helix and E-F loops colored red and the C-D loop represented as black sticks. For clarity the remainder of the m04 molecule is colored gray.

**FIGURE 2. The m04 molecular surface.** A, positive (blue) and negative (red) electrostatic potentials are shown on the surface of m04. Sugar moieties are represented as yellow spheres and are numbered according to the asparagine residue to which they are attached. B, surface view of m04 (gray) in the same orientation as A. The three surface-exposed hydrophobic residues (Ile\(^{90}\), Leu\(^{90}\), and Phe\(^{90}\)) are displayed as sticks and colored green.
determining region 1 of the Ig-V domain. A hydrophobic patch comprising Ile\(^90\), Leu\(^91\), and Phe\(^93\) is solvent-exposed on the outer face of the AA/H11032IDEE/H11032F sheet (Fig. 2B). Such surface-exposed hydrophobic residues are often considered to be involved in protein-protein interactions (22).

Structural Comparisons—The results of a Dali search (23) to identify proteins of similar structure to m04 is highly populated by Ig-V domain-containing proteins such as antibodies, T-cell receptor chains, and MHC-I molecules. The sequence conservation between m04 and Ig-V domain containing proteins is typically low (the top 100 Dali results range from 6 to 15% sequence identity). The top Dali “hit” was the variable domain from the heavy chain of an antibody Fab fragment (Protein Data Bank code 2W9D), which has a Z-score of 4.6 and possesses an root mean square deviation of 3.7 Å over 85 aligned C\(^\alpha\) atoms. Although the \(\beta\)-strand topology is similar, only the A’, C, D, and I strands of m04 superpose well, with the remainder of the secondary structure elements being tilted, translated, or absent (Fig. 3A). Moreover, m04 possesses a number of atypical features that are not normally associated with Ig-V domains. In particular, the N-terminal A/A’ strands occupy a position normally associated with the G-strand in an Ig-V domain, a topology that is unique to m04. This association is strengthened by an additional disulfide bond connecting Cys\(^{29}\) of the A’ strand to Cys\(^{146}\) of the I strand. Another peculiar feature of the A’ strand is that it leads almost immediately into a truncated B strand on the opposing sheet. Such an arrangement is reminiscent of that found in the H-type Ig fold where the D strand is a direct continuation of the C’ strand (24). In addition to the termini, the length and nature of the interconnecting loops deviates considerably from those found in the classical antibody Ig-V domains. For the most part, these loops are extended, as is evident in the large spacing (98 residues) between Cys\(^{46}\) and Cys\(^{141}\) that make up the canonical disulfide bridge (Fig. 3B). This distance is considerably larger than the 63–76 residues that are typically found within the disulfide bridge of Ig-V domains (25). Much of the additional length is derived from the E-F and F-G loops, the latter of which deviates around the molecule to pack on top of the short E-strand. These two loops are the most flexible regions of the structure with average side chain B-factors of 119 (residues 93–106) and 110 (residues 114–119), respectively. Thus, m04 possesses a number of unique features in comparison with other members of the Ig-V superfamily.

A Conserved Hydrophobic Interior—The m02 family members contain a number of conserved motifs that include a CXLXXC(L/P)(W/R)o motif found in m02–m06 and a NAWXXX(E/H)WO motif found in m03–m06. Our structure permitted a detailed understanding of the role of these motifs, as well as providing insight into the structure of other m02 family members. Surprisingly, despite being well separated in the m04 primary structure, the CXLXXC(L/P)(W/R)o and NAWXXX(E/H)WO motifs are intimately associated in the tertiary structure. Although the former occupies the majority of the I strand, the latter encompasses the entire length of the

![FIGURE 3. Structural conservation between m04 and an Ig-V domain. A, the structure of m04 was aligned with the highest scoring Dali match (labeled Fab 2W9D (46)). Both structures are shown side by side and overlaid. B, structure-based alignment of m04 with the best five Dali matches. Secondary structure elements are colored as in Fig. 1. The m04 strand assignment is shown above the sequences. Cysteine residues that form the canonical Ig-V disulfide bond and the analogous residues in m04 are displayed as green sticks (in A) and underlined (in B).](https://example.com/figure3.png)
adjacent D strand (Fig. 4A). Of particular prominence are the two Trp residues present in the NAXWX(E/H)Wo motif. Each of these residues assembles a myriad of hydrophobic residues to form two large but separate clusters that lie at the core of the Ig domain and mini platform regions. The first of these, Trp57 is located near the base of the D-strand and projects upwards toward the C-D loop (Fig. 4B). In doing so, Trp57 makes extensive van der Waals interactions with Leu143 and Met145 from the CXLXXC(L/P)(W/R)o motif, as well as Leu48, Pro49, and Val51 from the C-D loop, Tyr76 and Ser78 from the D-strand, and Pro84 from the D-E loop. The second tryptophan (Trp61) is located near the N terminus of the D-strand and is similarly surrounded by an array of hydrophobic side chains emanating from no less than seven separate β-strands (Fig. 4C) that include the C- (Leu44), D- (Leu59), E- (Val79), G- (Val113), H- (Leu125) and I-strands (Val139). A lone hydrogen bond to Ser74 is the only side chain-derived polar contact of note in the vicinity. Thus, these two motifs act to stabilize the three-dimensional scaffold of m04.

A Prototype for the m02 Superfamily—The m04 structure now provides an opportunity to establish a structural correlation between m02 family members. Although m02–m06 form a more closely related subclass within the m02 superfamily, m07–m16 also possess some related sequence motifs. Based on an analysis of the sequence of m02–m16 from MCMV64, we suggest that all family members with the exception of m12 and m13 are likely to adopt an m04-like fold. This proposition is based on three observations. First, both cysteine residues of the CXLXXC(L/P)(W/R)o motif are highly conserved (Fig. 5). Moreover, their binding partners that are involved in formation of the canonical (Cys24) and unique (Cys45) disulfide bonds are present in all sequences with the exception of m03 (Fig. 5). This correlation suggests that the majority of the m02 family members form not only an Ig-V-like domain, but also one with the unusual A’-strand positioning observed in m04. Second, the two prominent Trp57 and Trp61 residues from the NAXWX(E/H)Wo motif and the hydrophobic cluster within which they are enshrined (blue letters in Fig. 5) are well conserved throughout the family, suggesting they share a similar core packing arrangement to that observed in m04. Finally, most of the variability in sequence length is found in regions that would be predicted to be flexible loops (Fig. 3B), suggesting that the core secondary structure is, for the most part, con-

FIGURE 4. Molecular interactions within the m04 hydrophobic core. A, the side chains of conserved residues of the CXLXXC(L/P)(W/R)o and NAXWX(E/H)Wo motifs, which are located on the I-strand (cyan) and D-strand (purple), respectively, are shown as sticks. The main chain atoms of nonconserved residues of the D- and I-strands are also shown as sticks. B and C, close-up views of the clusters of hydrophobic residues surrounding Trp57 (B) and Trp61 (C) of the NAXWX(E/H)Wo motif. Hydrogen bonds between side chain atoms are shown as dashed lines.
served. Thus, it appears that most members of the m02 superfamily will adopt an Ig-V-like fold.

Direct Binding of m04 to MHC-I—To investigate the interaction between m04 and MHC-I, we performed SPR measurements on the soluble extracellular domains expressed in HEK 293 and E. coli, respectively. Although previous reports suggested that the m04-MHC-I interaction resides within their respective transmembrane domains, we observed a direct interaction between the extracellular domain of m04Smith with the murine MHC-I molecule H2-Dk (Fig. 6A). The affinity of the interaction was 52.0 ± 5.0 μM, well inside the typical range observed in immune recognition. Next we assessed the role of bound peptide and origin of the β2m subunit on m04 binding. The presence of a 9-mer peptide (RL9) or substitution of murine with human β2m had little impact on the binding affinity of m04Smith to H2-Dk, with affinity values of 71.8 ± 4.3 and 85.5 ± 4.3 μM, respectively (Fig. 6B and Table 2). Importantly, m04Smith did not bind appreciably to the human MHC-I molecule HLA-B*5701 (Fig. 6A), indicating that m04 is specific for murine MHC-I. Finally, we expanded the scope of our binding analysis to include m04 glycoproteins from two additional MCMV isolates, as well as two further MHC-I molecules.

FIGURE 5. Multiple sequence alignment of the m02 superfamily. The predicted extracellular domains of m02–m16 from MCMV G4 strain were aligned using ClustalW2. Because MCMV<sup>str</sup> does not possess a gene encoding m12, the m12 sequence shown is derived from the MCMV<sup>Smith</sup>. The secondary structure of m04 is shown above the sequences. Conserved cysteine residues involved in disulfide bonding in m04 are highlighted in red. Conserved hydrophobic residues that interact with Typ<sup>57</sup> and Typ<sup>61</sup> at the core of the m04 fold are highlighted in blue. Potential N-linked glycosylation sites are underlined. The CXLXCL(P/W)Ro and NA<sub>X</sub>WXX(E/H)Wo motifs are boxed. The m04 secondary structure is displayed above the sequences.
FIGURE 6. The m04-MHC-I interaction. SPR sensograms (top panels) and equilibrium binding curves (bottom panels) are shown for the binding of various m04 variants to a number of MHC-I molecules (A). B, the importance of the origin of the β2m subunit (murine (m) or human (h)) and the bound peptide was assessed. ND, not determined.
results of this analysis are summarized in Table 2. Whereas all of the m04s tested bound poorly or not at all to the human HLA-B*5701, all of m04Smith, m04G4, and m04WR211 bound to H2-Ld, Dd, and Dk. Overall the affinity of the interactions was relatively low, with the highest being the m04Smith-H2-Dk interaction (52.0 ± 5.0 μM) and the lowest being the m04WR211-H2-Dk interaction (478.6 ± 51.7 μM) (Fig. 6A). In general, m04Smith and m04G4, which share 93.6% sequence identity, bound to the panel of MHC-I molecules similarly and with higher affinity than the more distantly related m04WR211, which shares only 57.6% identity with m04Smith. Of the MHC-I molecules tested, H2-Dd displayed the weakest binding to the various m04 variants (K_d > 200 μM). Overall, our data suggest that m04 is capable of binding a broad range of murine MHC-I molecules with low to moderate affinity.

**DISCUSSION**

The down-regulation of MHC-I on the surface of infected cells is a common strategy employed by viruses to avoid detection by cytotoxic T cells. Accordingly, many viruses encode molecules that bind to MHC-I in the ER and either prevent its egress to the cell surface or redirect it to the lysosome for degradation. In MCMV, these functions are performed by m152, an MHC-I like molecule of the m145 family, and m06, whose structure has not yet been characterized (26). However, reduced levels of MHC-I on the surface of MCMV-infected cells results in a loss of signaling from inhibitory Ly49 receptors, thereby triggering NK cell activation (2). Perhaps to provide a balance between the evasion of NK and T cell responses, MCMV also encodes an additional glycoprotein, m04, whose role is to shuttle MHC-I to the cell surface (14). This strategy may have been counteracted, however, because NK cells have also evolved activating Ly49 receptors that are capable of recognizing infected cells of certain H-2 haplotypes in the presence of virally encoded m04 (15, 16).

As a first step toward understanding how Ly49 receptors recognize MHC-I in an m04-specific manner, we determined the crystal structure of m04 to 3Å resolution and investigated its interaction against a panel of pMHC-I molecules. Surprisingly, despite sharing little amino acid sequence identity, m04 possesses an Ig-V like β-sandwich fold that is decorated with a number of additional features including several extended loops and a prominent α-helix. Such insertions within this scaffold has previously been observed and in some examples even extra domains (in the case of hemocyanin (27), nuclear factor κB (28), and cytochrome f (29)) can be accommodated (30). In the case of m04, the presence of the α4 helix, which lies parallel to the C-D and E-F loops and is supported by the AA IDEE′F sheet gives the part of the molecule distal to the Ig domain a mini platform-like appearance. These regions could be interpreted as analogous to the α1 helix, peptide, α2 helix, and β-sheet scaffold of the MHC-I platform respectively, although in m04 the “peptide” and “helices” run parallel to the β-strands of the platform rather than diagonal as is the case in MHC-I. Whether this dual nature (Ig and platform) of m04 is related to its potential to interact with both MHC-I and Ly49 receptors is unknown.

The structure of m04 is the first documented example of an MCMV immunoevasin that adopts an Ig-like fold. However, in addition to MHC-I homologs such as UL18 (31) and UL142 (32), human CMV also encodes a number of Ig-like immune-modulating proteins that include UL141 and a cluster of genes located within the unique short (US) region (33, 34). Although they are unrelated in amino acid sequence, a number of parallels can be drawn between the unique short gene cluster and the m02 family. Indeed US2 and US11 are functionally similar to m06 in that they bind to newly assembled MHC-I molecules in the ER and target them for proteosomal degradation (35, 36).

It is interesting that both human and murine CMV appear to have independently evolved both MHC-I like and Ig-like immunoevasins. Indeed, the production of homologs of host genes that function in normal immune responses appears to be a common strategy employed by viruses to subvert the host.
immune system. Other such examples include the human CMV encoded UL146 (chemokine), UL33 and UL78 (chemokine receptors), and UL144 (TNF receptor) (37). These observations imply that the mechanism by which these immunoevasins act involves molecular mimicry, an interpretation that is supported by structural studies into UL18, which binds host LIR-1 in an almost identical manner to MHC-I (31). However, more recent studies suggest that virally encoded host homologs can mediate their function using novel unanticipated binding modes. For example, although possessing an MHC-I like structure, m157 does not bind to the Ly49 lectin-like domain in a manner analogous to MHC-I. Instead, m157 targets an aromatic peg motif within the helical Ly49 stalk regions (38, 39). Moreover, UL141, an Ig-like human CMV immunoevasin interacts with TRAIL receptor 2 in a unique fashion, resulting in its attenuated cell surface expression (34).

Such studies highlight the versatility of the MHC-I and Ig folds in mediating protein-protein interactions. Indeed, the three Ig-like proteins whose structures have been determined in complex with MHC-I all employ distinct binding modes that are tailored to suit their particular function. Killer immunoglobulin receptor receptors for example, which possess two or three tandem Ig domains, bind to MHC-I on top of the platform at a site that permits a high degree of allelic specificity (40, 41). In contrast, CD8 and US2 bind to more conserved regions away from the peptide-binding platform, presumably to permit binding to a wide array of MHC-I alleles. Although the m04-MHC-I binding site remains unknown, we suggest that it is likely to be similar to the site employed by US2 or CD8. This contention is supported by our structural and binding data. Initially, m04 adopts a compact structure with only a relatively short stalk region. Accordingly, because m04 binds to MHC-I in cis (within the plane of the same membrane), it is unlikely to be able to extend far enough away from the membrane to contact the top of the peptide-binding platform. This leaves the more conserved regions of the MHC-I molecule, including the underside of the platform, as well as the α3 domain and β2m subunit. Such a site is in agreement with our binding data, which indicates that m04 binds a wide range of murine MHC-I molecules, with large variations in the m04 sequence resulting in only subtle differences in affinity. Although we did not see significant differences in the affinity of m04 for MHC-I coupled to human versus murine β2m, we do not discount the possibility that the β2m forms part of the m04 binding site on account of the high level of sequence conservation of this subunit.

The m04 ectodomain sequence is highly variable among MCMV isolates (17), presumably because of selective pressure exerted by activating Ly49 receptors. Indeed, in the case of m157, whose sequence is also highly variable (42), infection of MCMV-resistant mice results in the rapid emergence of escape mutants (43, 44). Accordingly, mapping sites of variability in highly related m04 sequences may provide clues as to the location of molecular surfaces that are subject to immune selective pressure. The sequences of the m04<sup>Smith</sup> and m04<sup>K183</sup> ectodomains are highly similar to that of m04<sup>Δ4</sup>, containing only 17- and 26-amino acid differences, respectively (17). When mapped on to the m04<sup>Δ4</sup> structure, the vast majority of these differences are located in surface-exposed residues (Fig. 7), indicating that they likely impact on protein-protein interactions rather than disrupting the m04 fold. In particular, these variable residues cluster to a single face of the molecule that includes the B-C, D-E, and H-I loops and the top of the mini platform (Fig. 7A). In contrast, the opposing face of the molecule that includes the outer face of the AA’ IDE’F sheet and the N-linked glycan moieties appears to be more conserved (Fig. 7B).

By aligning the sequences of the m02 family members based on the structure of m04, we suggest that m02–m11 and m14–m16 are likely to adopt an Ig-like fold. Our exclusion of m12 and m13 from this grouping is primarily based on the absence of the conserved cysteine residues in these members. However, it is noteworthy that there are now many examples of Ig domains that lack the canonical disulfide bond (45), so it is not inconceivable that m12 and m13 may also possess Ig-related folds. Accordingly, it is possible that the m02 family arose by duplication from a single ancestral gene that encoded an Ig-like protein. As such, m04 represents a prototype for understanding the structure and function of this large immunoevasin family.

Acknowledgments—We thank Anthony Scalzo for the kind gift of the m04 DNA templates, as well as the Monash Macromolecular Crystalisation Facility and the staff at the Australian Synchrotron MX2 Beamline for expert assistance.

REFERENCES

1. Petersen, J. L., Morris, C. R., and Solheim, J. C. (2003) Virus evasion of MHC class I molecule presentation. J. Immunol. 171, 4473–4478
2. Karlhofe, F. M., Ribaudo, R. K., and Yokoyama, W. M. (1992) MHC class I allointagen specificity of Ly-49+ IL-2-activated natural killer cells. Nature 358, 66–70
3. Berry, R., Rossjohn, J., and Brooks, A. G. (2014) The Ly49 natural killer cell receptors: a versatile tool for viral self-discrimination. Immunol. Cell Biol. 92, 214–220
4. Reddehase, M. J. (2002) Antigens and immunoevasins: opponents in cytomegalovirus immune surveillance. Nat. Rev. Immunol. 2, 831–844
5. Sissons, J. G., Carmichael, A. J., McKinney, N., Sinclair, J. H., and Wills, M. R. (2002) Human cytomegalovirus and immunopathology. Springer Semin. Immunopathol. 24, 169–185
6. Dunn, W., Chou, C., Li, H., Rai, H., Patterson, D., Stolc, V., Zhu, H., and Liu, F. (2003) Functional profiling of a human cytomegalovirus genome. Proc. Natl. Acad. Sci. U.S.A. 100, 14223–14228
7. Rawlinson, W. D., Farrell, H. E., and Barrell, B. G. (1996) Analysis of the complete DNA sequence of murine cytomegalovirus. J. Virol. 70, 8833–8849
8. Adams, E. J., Joo, Z. S., Venook, R. T., Boulanger, M. J., Arase, H., Lanier, L. L., and Garcia, K. C. (2007) Structural elucidation of the m157 mouse cytomegalovirus ligand for Ly49 natural killer cell receptors. Proc. Natl. Acad. Sci. U.S.A. 104, 10128–10133
9. Mans, J., Natarajan, K., Balbo, A., Schuck, P., Eidel, D., Hess, S., Robinson, H., Simic, H., Jonjie, S., Tiemessen, C. T., and Margulies, D. H. (2007) Cellular expression and crystal structure of the murine cytomegalovirus major histocompatibility complex class I-like glycoprotein, m153. J. Biol. Chem. 282, 35247–35258
10. Natarajan, K., Hicks, A., Mans, J., Robinson, H., Guan, R., Maruiza, R. A., and Margulies, D. H. (2006) Crystal structure of the murine cytomegalovirus MHC-I homolog m144. J. Mol. Biol. 358, 157–171
11. Wang, R., Natarajan, K., Revilla, M. J., Boyd, L. F., Zhi, L., Zhao, H., Robinson, H., and Margulies, D. H. (2012) Structural basis of mouse cytomegalovirus m152/gp40 interaction with RAE1γ reveals a paradigm for MHC/MHC interaction in immune evasion. Proc. Natl. Acad. Sci. U.S.A. 109, E3578–E3587
12. Oliveira, S. A., Park, S. H., Lee, P., Bendelac, A., and Shenk, T. E. (2002) Murine cytomegalovirus m02 gene family protects against natural killer cell-mediated immune surveillance. J. Virol. 76, 885–894
13. Reusch, U., Murnani, W., Lucin, P., Burgert, H. G., Hengel, H., and Koszinowski, U. H. (1999) A cytomegalovirus glycoprotein re-routes MHC class I complexes to lysosomes for degradation. EMBO J. 18, 1081–1091
14. Kleijn, M. F., Huppa, J. B., Lucin, P., Mukherjee, S., Farrell, H., Campell, A. E., Koszinowski, U. H., Hill, A. B., and Ploegh, H. L. (1997) A mouse cytomegalovirus glycoprotein, gp34, forms a complex with folded class I MHC molecules in the ER which is not retained but is transported to the cell surface. EMBO J. 16, 685–694
15. Kiedczewska, A., Pyzik, M., Sun, T., Krmptom, A., Lodoen, M. B., Munks, M. W., Babic, M., Hill, A. B., Koszinowski, U. H., Joniec, S., Janier, L. L., and Vidal, S. M. (2009) Ly49P recognition of cytomegalovirus-infected cells expressing H-2Dk and CMV-encoded m04 correlates with the NK cell antiviral response. J. Exp. Med. 206, 515–523
16. Pyzik, M., Charbonneau, B., Gendron-Fonthiend, E. M., Babić, M., Krmptom, A., Joniec, S., and Vidal, S. M. (2011) Distinct MHC class I-dependent NK cell-activating receptors control cytomegalovirus infection in different mouse strains. J. Exp. Med. 208, 1105–1117
17. Corbett, A. J., Forbes, C. A., Moro, D., and Scalzo, A. A. (2007) Extensive sequence variation exists among isolates of murine cytomegalovirus within members of the m02 family of genes. J. Gen. Virol. 88, 758–769
18. Aricescu, A. R., Lu, W., and Jones, E. Y. (2006) A time- and cost-efficient system for high-level protein production in mammalian cells. Acta Crystallogr. D Biol. Crystallogr. 62, 1243–1250
19. Clements, C. S., Kjer-Nielsen, L., Kostenko, L., McCluskey, J., and Rossjohn, J. (2006) The production, purification and crystallization of a soluble form of the nonclassical MHC HLA-G: the essential role of coalt. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 62, 70–73
20. Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H., and Hung, L. W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. Acta Crystallogr. D Biol. Crystallogr. 65, 582–601
21. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674
22. Pang, S. S., Berry, R., Chen, Z., Kjer-Nielsen, L., Perugini, M. A., King, G. F. Wang, C., Chew, S. H., La Gruta, N. L., Williams, N. K., Beddow, T., Tigañas, T., Cowieson, N. P., Godfrey, D. I., Purcell, A. W., Wilce, M. C., McCluskey, J., and Rossjohn, J. (2010) The structural basis for autonomous dimerization of the pre-T-cell antigen receptor. Nature 467, 844–848
23. Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. Nucleic Acids Res. 38, W545–W549
24. Bork, P., Holm, L., and Sander, C. (1994) The immunoglobulin fold. Structural classification, sequence patterns and common core. J. Mol. Biol. 242, 309–320
25. Smith, D. K., and Xue, H. (1997) Sequence profiles of immunoglobulin and immunoglobulin-like domains. J. Mol. Biol. 274, 530–545
26. Holttappes, R., Gillert-Marien, D., Thomas, D., Podlech, J., Deegen, P., Herter, S., Oehrlein-Karpi, S. A., Strand, D., Wagner, M., and Reddehase, M. J. (2006) Cytomegalovirus encodes a positive regulator of antigen presentation. J. Virol. 80, 7613–7624
27. Volbeda, A., and Hol, W. G. (1989) Crystal structure of hexameric haemocyanin from Panulirus interruptus refined at 3.2 A resolution. J. Mol. Biol. 209, 249–279
28. Müller, C. W., Rey, F. A., Sodeoka, M., Verdin, G. L., and Harrison, S. C. (1995) Structure of the NF-κB p50 homodimer bound to DNA. Nature 373, 311–317
29. Martinez, S. E., Huang, D., Szczepaniak, A., Cramer, W. A., and Smith, J. L. (1994) Crystal structure of chloroplast cytochrome f reveals a novel cytochrome fold and unexpected heme ligation. Structure 2, 95–105
30. Halaby, D. M., Poupon, A., and Mornon, J. (1999) The immunoglobulin fold family: sequence analysis and 3D structure comparisons. Protein Eng.