Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

The structure of cytomegalovirus immune modulator UL141 highlights structural Ig-fold versatility for receptor binding

Ivana Nemčovičová[a,b] and Dirk M. Zajonc[a]*

[a]Division of Cell Biology, La Jolla Institute for Allergy and Immunology, 9420 Athena Circle, La Jolla, CA 92037, USA, and [b]Department of Molecular Medicine, Institute of Virology, Slovak Academy of Sciences, Dúbravská cesta 9, SK 84505 Bratislava, Slovakia

Correspondence e-mail: dzajonc@liai.org

Received 4 April 2013
Accepted 13 December 2013

PDB reference: UL141, 4jm0

Natural killer (NK) cells are critical components of the innate immune system as they rapidly detect and destroy infected cells. To avoid immune recognition and to allow long-term persistence in the host, *Human cytomegalovirus* (HCMV) has evolved a number of genes to evade or inhibit immune effector pathways. In particular, UL141 can inhibit cell-surface expression of both the NK cell-activating ligand CD155 as well as the TRAIL death receptors (TRAIL-R1 and TRAIL-R2). The crystal structure of unliganded HCMV UL141 refined to 3.25 Å resolution allowed analysis of its head-to-tail dimerization interface. A 'dimerization-deficient' mutant of UL141 (ddUL141) was further designed, which retained the ability to bind to TRAIL-R2 or CD155 while losing the ability to cross-link two receptor monomers. Structural comparison of unliganded UL141 with UL141 bound to TRAIL-R2 further identified a mobile loop that makes intimate contacts with TRAIL-R2 upon receptor engagement. Superposition of the Ig-like domain of UL141 on the CD155 ligand T-cell immunoreceptor with Ig and ITIM domains (TIGIT) revealed that UL141 can potentially engage CD155 similar to TIGIT by using the C′C′′ and GF loops. Further mutations in the TIGIT binding site of CD155 (Q63R and F128R) abrogated UL141 binding, suggesting that the Ig-like domain of UL141 is a viral mimic of TIGIT, as it targets the same binding site on CD155 using similar 'lock-and-key' interactions. Sequence alignment of the UL141 gene and its orthologues also showed conservation in this highly hydrophobic (L/A)X_6G 'lock' motif for CD155 binding as well as conservation of the TRAIL-R2 binding patches, suggesting that these host–receptor interactions are evolutionary conserved.

1. Introduction

Cytomegaloviruses (CMVs) are members of the subfamily *Betaherpesvirinae* of the family *Herpesviridae*. *Human cytomegalovirus* (HCMV) is a common cause of congenital viral infections and is a frequent opportunistic pathogen in transplant recipients and HIV patients (Alford *et al.*, 1990; Pass, 2005; Ho, 1991). The relative genomic complexity of HCMV is mirrored by its biological characteristics, as the most relevant cellular reservoirs of the latent virus and its sites for permissive replication have not been conclusively established and its pathogenesis is not well understood. Like other CMVs, HCMV has a very specific host range, but within a permissive host it enters and replicates in a wide variety of cell types (Ho *et al.*, 1991; Sinzger *et al.*, 1995).

The HCMV genome consists of 230–235 kb of double-stranded DNA and more than 160 predicted open reading frames (ORFs; Dolan *et al.*, 2004; Murphy *et al.*, 2003; Chee *et al.*, 1990; Davison *et al.*, 2003), while a recent report estimates

that HCMV can encode over 700 ORFs owing to the presence of alternate transcriptional start sites (Stern-Ginossar *et al.*, 2012). Virulent clinical isolates of HCMV have been shown to carry at least 19 additional genes (Cha *et al.*, 1996), designated UL133–UL151, that are contained in the unique long (UL) b′ region of the HCMV genome and many of which possess immunomodulatory function. The UL/b′ region is deleted in extensively passaged laboratory strains and its presence correlates with adverse effects in vaccinated persons, indicating that viral genes in this region play a significant role in controlling virulence. The recently characterized HCMV immunomodulatory protein UL141 is located in this UL/b′ region (Nemčovičová *et al.*, 2013; Smith *et al.*, 2013Tomasec *et al.*, 2005).

HCMV targets host receptors for viral entry into cells, while at the same time expressing immunomodulatory proteins, including viral ligands for members of the tumour necrosis factor (TNF) superfamily as well as poliovirus receptor-related (PRR) nectin-like molecules, to block host signalling pathways that would otherwise lead to immune recognition and viral clearance (Ware & Sedý, 2011; Yang & Bjorkman, 2008; Prod'homme *et al.*, 2010; Baumgarth *et al.*, 2008). Moreover, many of the targeted host proteins have basic cellular functions and are therefore required and even beneficial for the virus during the course of infection, while other host proteins could facilitate viral infection by performing functions that are unrelated to their normal role. TNF/TNFR members play important roles in regulating many biological functions, especially as prominent mediators of immune regulation, inflammatory responses, bone development and homeostasis (Cha *et al.*, 2000). It is known that TNF superfamily members play crucial roles in controlling herpesvirus infection by initiating the direct killing of infected cells and by enhancing immune responses (Cha *et al.*, 2000; Ware & Sedý, 2011). While TRAIL death receptors (TRAIL-DR) induce the apoptosis of herpesvirus-infected cells to maintain immune homeostasis, herpesvirus can in turn block apoptotic signalling and establish lifelong infection (Benedict *et al.*, 2003; Roy & Mocarski, 2007). UL141 has recently been shown to restrict cell-surface expression of TRAIL-DR (TRAIL-R1/DR4 and TRAIL-R2/DR5), while cells infected with an HCMVΔUL141 deletion strain were more susceptible to killing by TRAIL, suggesting that UL141 is required to inhibit the expression of both TRAIL death receptors (Smith *et al.*, 2013). However, NK cells express a multitude of different activating and inhibitory receptors, with the function of each cell being regulated by the integration of signals received from ligands presented on potential target cells (Lanier, 2008). As an example of viral immune evasive strategy, HCMV blocks T-cell activation by downregulating endogenous MHC-I (Ahn *et al.*, 1997; Furman *et al.*, 2002). Downregulation of host MHC-I would be noticed by a lack of engagement of inhibitory receptors on NK cells; however, the virus encodes its own MHC-I homologue (viral UL18) to activate the inhibitory receptor LIR-1/ILT-2 (Beck & Barrell, 1988; Chapman *et al.*, 1999; Prod'homme *et al.*, 2007). In addition, an HCMV-derived peptide acts to promote cell-surface expression of the nonclassical MHC-I molecule HLA-E, the ligand for the inhibitory receptor CD94 (Tomasec *et al.*, 2000; Ulbrecht *et al.*, 2000; Wang *et al.*, 2002). The activating receptor NKG2D recognizes eight ligands that are upregulated upon stress (infection), including MICA, MICB, ULBP1 and ULBP2. To block NK cell activation, the viral immune evasin UL16 can bind to and retain MICB, ULBP1 and ULBP2 in the endoplasmic reticulum (ER). Furthermore, viral UL112 targets the MICB transcript, while viral UL142 downregulates MICA (Chalupny *et al.*, 2006; Cosman *et al.*, 2001; Stern-Ginossar *et al.*, 2007; Wills *et al.*, 2005), thus abrogating cell-surface expression of NK cell-activating ligands and blocking NK cell activation.

Recently, it has been shown that HCMV targets CD112 (PRR-2; poliovirus receptor-related protein 2) and CD155 for proteasome-mediated degradation, thus reducing cell-surface expression of both NK cell-activating ligands for CD226/DNAM-1 during infection (Prod'homme *et al.*, 2010; Bottino *et al.*, 2003; Fuchs *et al.*, 2004; Tomasec *et al.*, 2005). More importantly, cell-surface expression of both CD112 and CD155 was restored when UL141 was deleted from the HCMV genome. While viral UL141 alone was found to be sufficient to mediate retention of CD155 in the ER, UL141 requires additional HCMV-encoded factors to downregulate the expression of CD112 (Tomasec *et al.*, 2005; Prod'homme *et al.*, 2010; Fuchs *et al.*, 2004; Bottino *et al.*, 2003).

Despite a strong host immune response, herpesviruses persist in a latent form, indicating the dynamic relationship between the host immune system and the virus, which results in a balance between host survival and viral control. Characterizing the structural and molecular bases of the interactions that occur between these HCMV proteins (such as viral UL141) and their cellular targets (such as TRAIL-related or PRR-related receptors) is crucial for our understanding of viral persistence and will ultimately facilitate vaccine and antiviral drug development.

Here, we report the crystal structure and binding-site characterization of UL141, a viral glycoprotein that has evolved to target two different NK cell signalling pathways by mimicking cellular ligands of both TNFRSF (TRAIL-DR) and Ig superfamily members (CD155).

2. Methods

2.1. Molecular cloning

The mature ectodomain of UL141 (amino acids 30–279, HCMV FIX strain, Gene ID 3077418) was PCR-amplified using the oligonucleotides 5′-CCGGGATCCCTCGTTCCCC-TTCGCCACCG-3′ and 5′-CCGGAATTCTCAGTGATGG-TGATGGTGATGTCCCCGAGTGGCCCAGGG-3′ and was cloned downstream of the gp67 signal sequence in the baculovirus transfer vector pAcGP67A (BD Biosciences) with the addition of a C-terminal hexahistidine tag. The same ectodomain of UL141 was also cloned into the mammalian expression vector pCR3.1 with the addition of a C-terminal Fc-fusion tag of human IgG_1. The ectodomain of human CD155

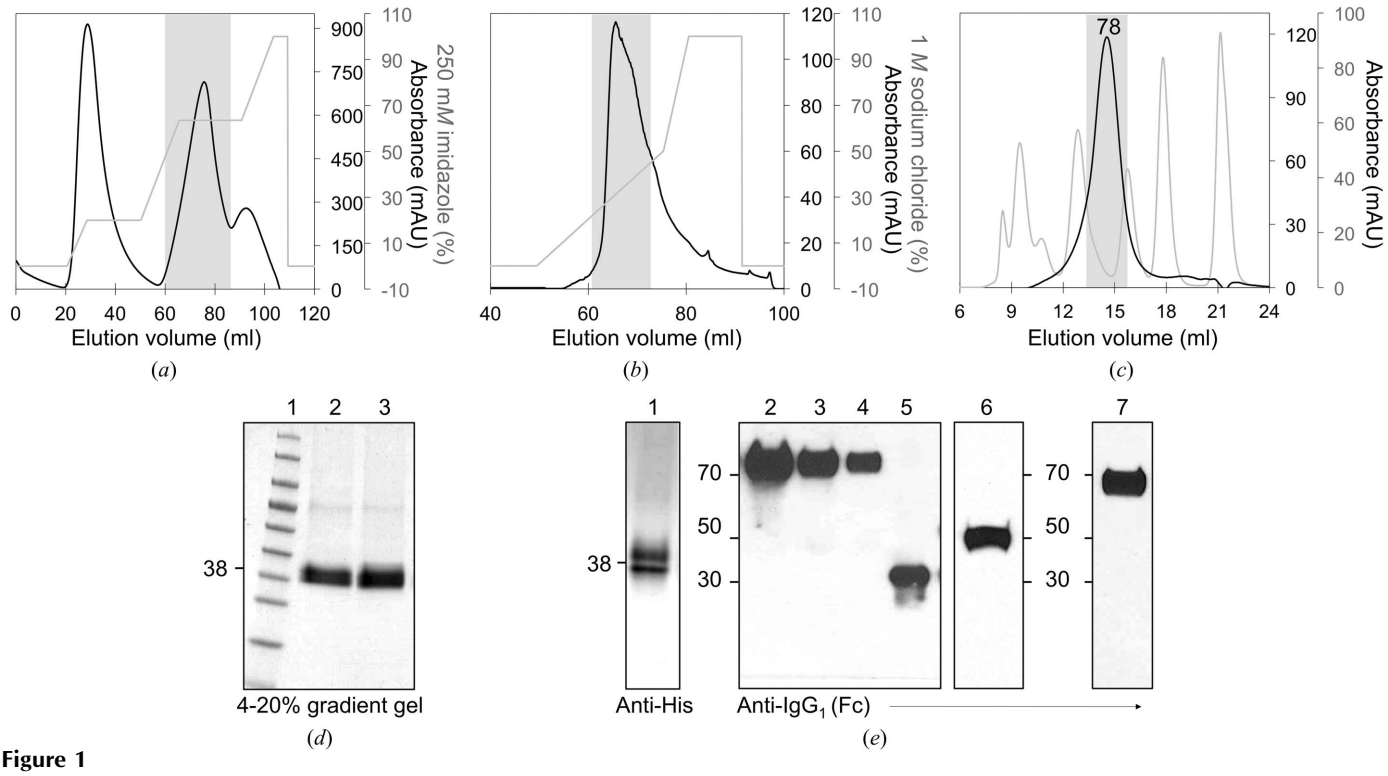

Figure 1
Purification of HCMV UL141 from Sf9 insect cells and Western blot analysis. (*a*) Metal-ion affinity chromatography using Ni–NTA agarose (HiTrap 1 ml column, GE Healthcare) and a linear/step gradient of 250 m*M* imidazole in 50 m*M* Tris–HCl pH 8.0. (*b*) Anion-exchange chromatography (Mono Q 1 ml column, GE Healthcare) performed using a linear gradient of 1 *M* sodium chloride. (*c*) Size-exclusion chromatography (Superdex S200 10/300 column, GE Healthcare). Shaded areas represent UL141-containing fractions that were pooled for subsequent purification steps. (*d*) 4–20% SDS–PAGE analysis of UL141: lane 1, protein marker (labelled in kDa); lane 2, reduced; lane 3, nonreduced. (*e*) Western blot analysis of His-tagged UL141 (lane 1), CD155-Fc-WT (lane 2), CD155-Fc-Q63R (lane 3), CD155-Fc-F128R (lane 4), Fc control (lane 5), TRAIL-R2-Fc (lane 6) and UL141-Fc (lane 7).

(amino acids 29–343, Gene ID 5817) was PCR-amplified using the oligonucleotides 5′-CCGGGATCCCGTCGTCGTGCAGGCGCCCACCCAGGTGCCCGGCTT-3′ and 5′-CCGCTGCAGTCAGTGATGGTGATGGTGATGGTTACGGGATATGCC-3′ and cloned downstream of the secretion-signal sequence in the mammalian expression vector pCR3.1 with and without the addition of a C-terminal Fc-fusion tag. The CD155 mutants Q63R and F128R were generated in the same vector using complementary pairs of single-stranded mutagenic PCR primers: 5′-ATGGAGGTGACGCATGTGTCACGGCTGACTTGGGCGCGGCATGGT-3′, 3′-ACCATGCCGCGCCCAAGTCAGCCGTGACACATGCGTCACCTCCAT-5′ and 5′-TACACCTGCCTGTTCGTCACGCGCCCGCAGGGCAGCAGGAGCGTG-3′, 3′-CACGCTCCTGCTGCCCTGCGGGCGCGTGACGAACAGGCAGGTGTA-5′, respectively. The UL141 dimerization-deficient triple mutation (N46A, E61A, T71F) was generated using the QuikChange II Multi-site Mutagenesis Kit (Stratagene, La Jolla, California, USA) using the single-stranded oligonucleotides 5′-CCGAAAAGATGTGGGCCGAGGCTTATGAGACCACGTCGCCGGCG-3′, 5′-CCGGTGTTGGTCGCCGAGGGAGCGCAAGTTACCATCCCCTGCACG-3′ and 5′-ACCATCCCCTGCACGGTCATGTTCCACTCCTGGCCCATGGTCTCC-3′. The ectodomain of human TRAIL-R2 (amino acids 58–184, Gene ID 8795) was cloned as an Fc-fusion protein into a mammalian expression system as described previously (Nemčovičová *et al.*, 2013). The correct sequences of the inserted genes and successful mutagenesis were confirmed by sequencing, and the corresponding plasmids were further amplified in bacteria (*Escherichia coli* XL10-Gold or *E. coli* JM109) using the Endofree Plasmid Maxi Kit (Qiagen, Valencia, California, USA) and maintained under sterile conditions.

2.2. Protein expression and purification

UL141 was expressed using the baculovirus expression system in *Spodoptera frugiperda* Sf9 cells (see Supporting Information §S1[1]). After 84 h of expression in insect-cell medium at 301.15 K, Sf9 cells and debris were removed from the UL141-containing culture supernatant by centrifugation (5 min at 1000*g* followed by an additional 10 min at 6000*g*). The supernatant was concentrated to 300 ml and gradually exchanged against 50 m*M* Tris–HCl pH 8, 300 m*M* NaCl, 20 m*M* imidazole by tangential flowthrough filtration using 10 kDa molecular-weight cutoff membranes (Millipore Pelicon 2 filtration device). The UL141 was purified by Ni^{2+}-affinity chromatography (Fig. 1*a*) using an imidazole step gradient (HisTrap 1 ml column, GE Healthcare). Next, UL141-containing fractions were pooled and dialyzed at

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: XB5073).

277.15 K against 10 m*M* Tris–HCl pH 8 buffer for subsequent purification by anion-exchange chromatography using Mono Q (GE Healthcare) and a linear gradient of 0–1 *M* sodium chloride (Fig. 1*b*). The UL141-containing fractions were pooled, concentrated to 100 µl and injected onto a Superdex S20 (GE Healthcare) size-exclusion chromatography (SEC; Fig. 1*c*) column. Peak fractions with greater than 95% purity as shown by SDS–PAGE (Fig. 1*d*) were pooled, concentrated and diluted twofold with pure water to reach a final buffer concentration of 25 m*M* HEPES pH 7.5, 75 m*M* NaCl. The protein migrated as a major band at 38–40 kDa on both reducing and nonreducing SDS–PAGE (Fig. 1*d*, lanes 2 and 3) as well as on a Western blot (Fig. 1*e*, lane 1).

Fc-fusion proteins (UL141-Fc, CD155-Fc and TRAIL-R2-Fc) were produced in a mammalian expression system in 293T cells. Cells were grown at 310.15 K and 5% CO_2 in VLE Dulbecco's MEM (DMEM; Biochrome AG, Berlin) supplemented with 10%(*v*/*v*) fetal calf serum (FCS), 2 m*M* L-glutamine and 2 mg ml^{-1} gentamicin (Sandoz Pharmaceuticals). Transfected 293T cells were maintained in serum-free DMEM during the transfection process. Plasmid DNAs were prepared using the Endofree Plasmid Maxi Kit (Qiagen, Valencia, California, USA) and maintained under sterile conditions. The confluent 293T cells were passaged in T-175 flasks in serum-rich DMEM and incubated at 310 K with 5% CO_2. As a detaching component, 0.05% trypsin–EDTA solution was used to further maintain the cells. 293T cells were transfected using the Xfect transfection reagent (Clontech) and were subsequently maintained for 24, 48 and 72 h. The transfection mixture containing 100 µl Xfect reaction buffer, 5 µg DNA (mixed together) and 1.5 µl Xfect polymer (for one T-25 flask) was incubated for 10 min at room temperature and transferred dropwise to seeded 293T cells in a T-25 flask containing 3 ml serum-free DMEM medium. After 4 h of transfection, the medium was changed to 3 ml serum-rich DMEM medium containing antibiotics and L-glutamine. After 24 h of expression, the medium was replaced by fresh medium and the supernatant was collected for harvesting, while expression of the rest of the cells in fresh medium continued for a further 24 or 48 h. Freshly expressed proteins were run on SDS gradient polyacrylamide gels and transferred to nitrocellulose membranes. The blots were probed with anti-human IgG-HRP conjugate antibodies (Bio-Rad) for Fc-fusion proteins or with mouse anti-penta-His conjugate and anti-mouse IgG-HRP conjugate antibodies (Sigma).

The expression levels of Fc-fused proteins are shown by Western blots (Fig. 1*e*, right) and these proteins were used directly from the culture supernatant for surface plasmon resonance (SPR) studies.

2.3. Crystallization

Initial crystallization trials were carried out by robotic crystallization (Phoenix, Art Robbins Instruments) using the sitting-drop vapour-diffusion method at room temperature (~295 K) as well as 277 K in a three-well INTELLI-PLATE 96 (Art Robbins Instruments). Over 1000 conditions were screened using a variety of commercially available sparse-matrix crystallization screens, finally leading to only one condition that yielded a crystal suitable for X-ray diffraction measurement. This crystal (Supplementary Fig. S1*a*) could not be reproduced and subsequent trials yielded crystals with negligible diffraction power (Supplementary Figs. S1*b* and S1*c*). A single UL141 crystal with approximate dimensions of 60 × 30 × 30 µm was grown by mixing 0.2 µl protein solution (4 mg ml^{-1}) with 0.2 µl reservoir solution [0.2 *M* calcium acetate, 0.1 *M* imidazole pH 8, 10%(*w*/*v*) polyethylene glycol 8000] and incubating at 295 K for 3–4 d. The crystal was harvested with a nylon cryoloop and transferred into a 10 µl drop of 0.2 *M* calcium acetate, 0.1 *M* imidazole pH 8, 10%(*w*/*v*) polyethylene glycol 8000, 25%(*v*/*v*) glycerol. The crystal was soaked for at least 1 min before being cooled to 100 K by plunging into liquid nitrogen.

Table 1
Data-collection and refinement statistics for HCMV UL141.

Values in parentheses are for the highest resolution shell.

Crystal parameters	
Space group	$P3_221$
Unit-cell parameters (Å, °)	$a = b = 96.06$, $c = 136.07$, $\alpha = \beta = 90.00$, $\gamma = 120.00$
Data collection	
Resolution range (Å)	35.22–3.25 (3.33–3.25)
Wavelength (Å)	0.9698
R_{merge} (%)	9.30 (72.60)
$\langle I/\sigma(I)\rangle$	10.20 (3.10)
No. of unique reflections	10665 (1274)
Multiplicity	3.50 (3.50)
Completeness (%)	97.46 (99.46)
Refinement	
Resolution range (Å)	11.98–3.25
R factor (%)	20.14
R_{free} (%)	27.93
R.m.s. deviations	
Bond lengths (Å)	0.02
Bond angles (°)	2.35
No. of protein atoms	2863
No. of water molecules	0
Average *B* factor (Å^2)	42.11
Ramachandran analysis	
Favoured	85.40
Allowed	96.20
PDB code	4jm0

2.4. Data collection and processing

The UL141 crystal was tested remotely at the Stanford Synchrotron Radiation Laboratory (SSRL). Initial diffraction data were collected to 3.25 Å resolution on beamline 7-1. Owing to a space-group mismatch followed by incomplete data collection, the first data set alone was not amenable to structure determination. A complete data set from the same crystal was subsequently collected to 3.5 Å resolution using a MAR 325 CCD detector on SSRL beamline 9-2 with an oscillation angle of 1° and an exposure time of 15 s per image. Diffraction images from both data sets were merged and processed with *iMosflm* (Battye *et al.*, 2011), *POINTLESS* (Grosse-Kunstleve & Adams, 2002) and *SCALA* (Evans, 2006) within the *CCP*4 suite (Winn *et al.*, 2011) to a resolution of 3.25 Å. The crystal of HCMV UL141 belonged to the

trigonal space group $P3_221$. The unit-cell parameters were a = 96.06, b = 96.06, c = 136.07 Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. Data-collection and processing statistics for the merged data set used for structure determination are presented in Table 1.

2.5. Structure determination and refinement

Phasing and structure determination of HCMV UL141 was performed by the molecular-replacement (MR) method using *Phaser* (McCoy *et al.*, 2007) as part of the *CCP*4 suite (Winn *et al.*, 2011), using the protein coordinates from the recently determined crystal structure of UL141 in complex with TRAIL receptor 2 (TRAIL-R2; PDB entry 4ix9; Nemčovičová *et al.*, 2013) as a search model. This structure provided us with an independent model at higher resolution for the calculation of model phases of unbound native UL141 data at 3.25 Å resolution. One UL141 dimer per asymmetric unit was suggested by the resulting Matthews coefficient calculations (Matthews, 1968), with a V_M of 2.19 Å^3 Da^{-1} and a solvent content of 43%. The initial model was rebuilt using *ARP*/*wARP* (Cohen *et al.*, 2008) and underwent density modification using *Parrot* (Zhang *et al.*, 1997). After several rounds of iterative model building and refinement using *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010) and *REFMAC*5 (Murshudov *et al.*, 2011), the electron density improved and N-linked sugars were incorporated into the structure. Later rounds of refinement included TLS refinement (Painter & Merritt, 2006) with five groups per UL141 monomer. Intermolecular interactions were analyzed using the *Protein Interfaces, Surfaces and Assemblies* service (*PISA*; Krissinel & Henrick, 2007) at the European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html). TLS parameters were calculated using the *TLSMD* web server (http://skuld.bmsc.washington.edu/~tlsmd). Geometry was analyzed using *PROCHECK* (Laskowski *et al.*, 1993) and *MolProbity* (Chen *et al.*, 2010). R.m.s.d. values for the bound and unbound HCMV UL141 structures were calculated using the *SuperPose* (Maiti *et al.*, 2004) server. OMIT maps were generated using *SFCHECK* (Vaguine *et al.*, 1999) and the images of protein structures were prepared using *MacPyMOL* (v.1.5.0.4; Schrödinger).

2.6. Surface plasmon resonance (SPR) binding assay

After purification, the proteins were concentrated using Amicon centrifugal filter units (Millipore, 10K cutoff membrane) and the buffer was exchanged against 10 m*M* HEPES pH 7.4, 150 m*M* sodium chloride, 3 m*M* EDTA (Biacore running buffer). The proteins were diluted in Biacore running buffer containing 0.005% Tween 20 to appropriate concentrations prior to loading. Fc-fused proteins were used directly from cell-culture supernatant (see §2.2). SPR experiments were performed as reported previously (Nemčovičová *et al.*, 2013). Briefly, an anti-human Fc capture antibody was immobilized on a CM5 sensor chip (GE Healthcare) by amine coupling. In the experiments shown in Figs. 2(*a*) and 2(*c*) approximately 500 response units (RU) of the ligands TRAIL-R2 and CD155-Fc were captured on the sensor chip. The UL141 dimer (WT) and UL141 monomer (ddUL141 mutant 3) were prepared in running buffer at a concentration of

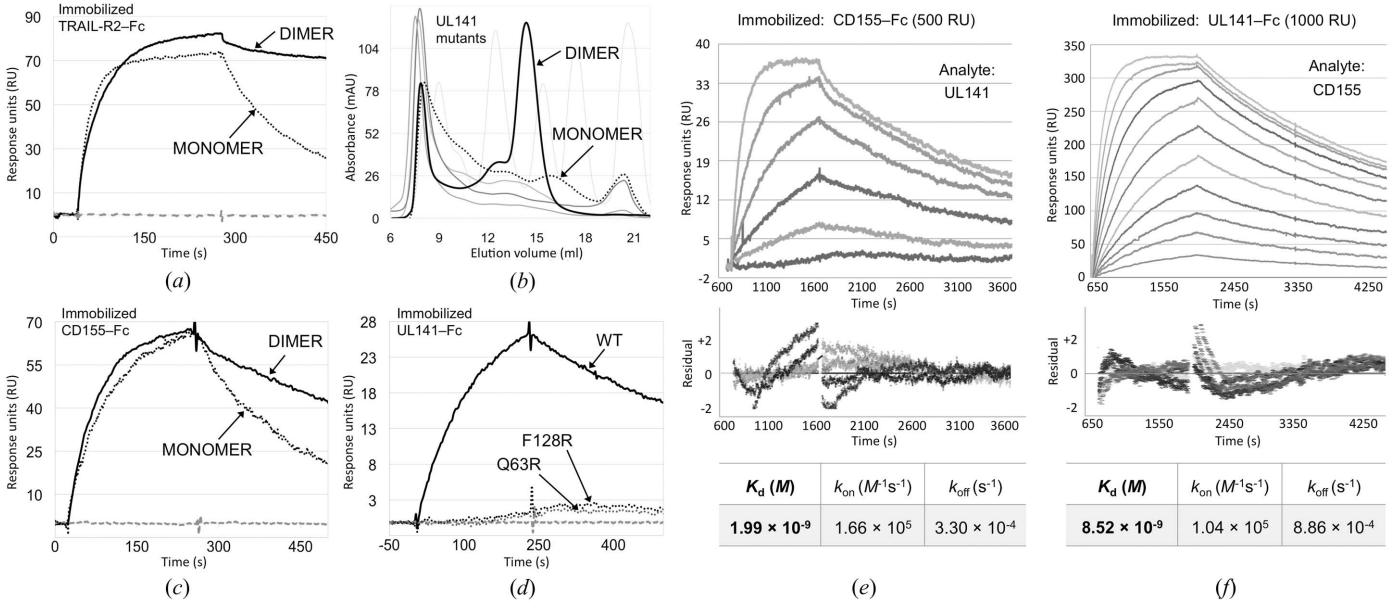

Figure 2
SPR binding analysis and SEC profiles. The SPR binding curves (*a*, *c*) for UL141 dimer (wild type, bold line) and UL141 monomer (ddUL141 mutant 3, dotted line) show that monomeric UL141 binds to immobilized TRAIL-R2 (*a*) or CD155 (*c*) with reduced avidity compared with wild-type UL141. (*b*) Size-exclusion elution profiles of UL141 mutants: mut 1, N46A; mut 2, E61A/T71F/L121W; dd mut 3, N46A/E61A/T71F; mut 4, N46A/E61A/T71F/L121W; mut 9, T71F/L121W. Most mutations lead to aggregation of UL141, while triple mutant 3 (N46A, E61A, T71F) shows a clear population of monomeric UL141 (dotted line). (*d*) The SPR binding responses for UL141 (wild type) show binding to CD155 (wild type, bold line) but not to CD155 mutants (Q63R or F128R, dotted lines). (*e*, *f*) Representative SPR trace for kinetic binding data for CD155-Fc *versus* UL141 according to Nemčovičová *et al.* (2013) (*e*) and UL141-Fc *versus* CD155 (*f*) including residual plots and statistics.

200 n*M*. These analytes were then injected for 5 min association, while dissociation was continued for 3–5 min, after which the chip was regenerated with a 30 s injection of 3 *M* $MgCl_2$ at 30 µl min^{-1}. Experiments were carried out at 291.15 K with a flow rate of 20 µl min^{-1}. As a negative control for nonspecific binding, human LTβR-Fc (lymphotoxin β receptor from the TNFR family) or human IgG_1 (Fc) was immobilized on the first flow channel. The response to the negative control (LTβR-Fc or Fc) and the buffer-only control as a background were subtracted from the final curves using the *BIAevaluation* software v.4.1. The experiment shown in Fig. 2(*d*) was performed similarly: approximately 200 RU each of CD155-Fc wild type (WT) and the two mutants Q63R and F128R were captured on the sensor chip. UL141 dimer (WT) was then injected as described above and the responses are shown. For comparison, the Biacore kinetics experiment shown in Fig. 2(*e*) was prepared as described previously (Nemčovičová *et al.*, 2013). In Fig. 2(*e*), the response of UL141 at different concentrations was monitored over CD155-Fc (∼500 RU) immobilized on the CM5 chip; *vice versa*, in Fig. 2(*f*) CD155 at different concentrations was monitored over UL141-Fc (∼1000 RU). The binding parameters (K_d, k_{on} and k_{off}) reported in these figures were calculated using the *BIAevaluation* software as mentioned above.

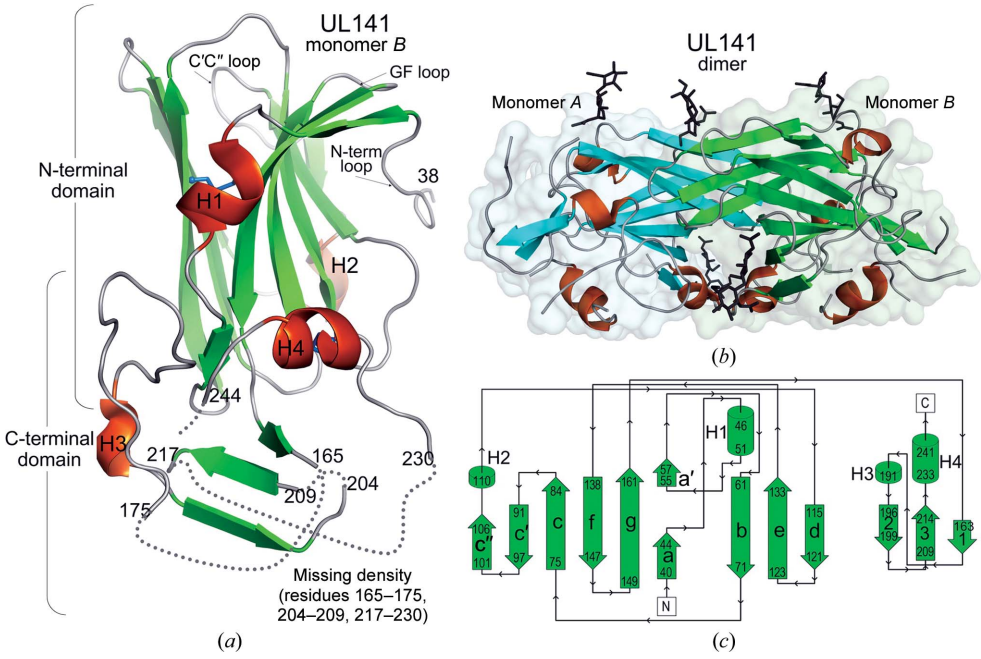

Figure 3
Structure of unliganded HCMV UL141. (*a*) Structure of the HCMV UL141 monomer (chain *B*) in ribbon representation with α-helices coloured orange and β-sheets shown in green. Two disulfide bonds, Cys67–Cys143 and Cys84–Cys234, are shown as blue sticks. (*b*) Cartoon representation of the HCMV UL141 dimer (monomers *A* and *B*). Monomers are coloured cyan and green, respectively, with the molecular surface shown in transparent colours. N-linked glycans are depicted as dark grey sticks. (*c*) Two-dimensional topology diagram generated by *PDBsum* (de Beer *et al.*, 2014; Laskowski, 2009) of UL141 monomer *B* with arrows from the N-terminus to the C-terminus.

3. Results and discussion

3.1. Structure determination of native HCMV UL141

For SPR and crystallographic studies, the UL141 ectodomain was expressed in *S. frugiperda* (Sf9) insect cells using a baculovirus-mediated expression system. The recombinant UL141 produced by this method gradually lost its receptor-binding activity within a week, suggesting that it was unstable in solution. Therefore, UL141 was freshly prepared and its ability to bind TRAIL-R2 or CD155 was assessed by surface plasmon resonance (SPR) immediately prior to crystallization. Active UL141 crystallized in the trigonal space group $P3_221$, with unit-cell parameters a = 96.06, b = 96.06, c = 136.07 Å, containing one UL141 dimer in the asymmetric unit. Two diffraction data sets were collected from the same crystal and were merged and processed to 3.25 Å resolution. The structure was determined by molecular replacement (MR) using the UL141 protein coordinates from the recently determined UL141–TRAIL-R2 complex as a search model (PDB entry 4i9x; Nemčovičová *et al.*, 2013). No density was observed in the model for the first eight N-terminal residues of the mature protein, suggesting that this region is disordered. Neither was density observed for the last 35 residues, which are likely to form a flexible tether connecting UL141 to the cell membrane. The disordered regions differ between the individual monomers of the UL141 dimer (see §3.2). The final crystal structure (Figs. 3*a* and 3*b*) was refined to a resolution of 3.25 Å with an R factor of 20.14% and an R_{free} of 27.93%.

3.2. UL141 structural conservation and predicted cysteine network

UL141 has no sequence similarity to any other known cellular protein and shares only weak homology with its family member UL14. The latter was identified by sequence alignment of human CMV UL141, its orthologues in chimpanzee, simian and rhesus CMV, and related UL14 genes (Supplementary Fig. S2). In addition, we further inspected and analyzed the conservation of the residues responsible for host–receptor binding across the species (Supplementary Fig. S2, residues in magenta). We identified eight out of 12 residues important in UL141 binding to TRAIL-R2 that are highly conserved in the viral sequences, suggesting that the UL141 host–receptor interaction is evolutionarily conserved across the species.

UL141 forms a noncovalent homodimer in solution (Figs. 1*c* and 1*d*) as well as in the crystal (Fig. 3*b*) (Nemčovičová *et al.*, 2013). Structural analysis of receptor-bound and unbound UL141 ectodomain clearly illustrates its head-to-tail arrangement formed by an N-terminal variable (V)-type immunoglobulin (Ig)-like domain followed by an additional C-terminal β-sheet domain (Figs. 3*a* and 3*b*). Inspection of the two-dimensional topology diagram of unliganded UL141 revealed shorter β-strands 1 and 3 of the C-terminal domain (Fig. 3*c*) owing to the high degree of disorder in this region. The unbound UL141 structure lacks clear electron density for residues 30–38, 165–175, 204–209, 217–230, 245–279 in monomer *A* and residues 30–32, 167–175, 203–209, 216–228, 244–279 in monomer *B*, as indicated in Fig. 3(*a*). Some of these residues were also missing in the structure of the UL141–TRAIL-R2 complex (residues 168–174, 199–207, 217–226 and 247–279 for both monomers), suggesting that these regions are either highly mobile or disordered owing to a lack of fully formed intramolecular disulfide bonds to stabilize the C-terminal domain.

The sequence of HCMV UL141 shows ten cysteines (Cys67, Cys84, Cys143, Cys173, Cys174, Cys200, Cys221, Cys226, Cys234 and Cys247) located in the extracellular part of the gene, eight of which are conserved within the viral sequences (Supplementary Fig. S2, green and yellow) and two of which are even conserved in common Ig domains (Supplementary Fig. S2, yellow). In the crystal structure, the Ig-like domain contains two disulfide bonds (Cys67–Cys143 and Cys84–Cys234), while Cys200 is unpaired and is the only ordered cysteine in the C-terminal appendix domain. However, we assume that the C-terminal appendix domain is not fully folded and that the five remaining cysteines (Cys173, Cys174, Cys221, Cys226 and Cys247) could also be involved in additional intramolecular disulfide bonds. If true, this would explain the partial instability and gradual loss of binding activity of recombinant UL141. Upon mapping the relative positions of the disordered cysteines, we postulate that the two cysteines (Cys221 and Cys226) in the disordered loop 217–230 are in close proximity to another two cysteines (Cys247 and Cys200, respectively) and could potentially form two additional disulfide bonds (Cys200–Cys226 and Cys221–Cys247). In particular, Cys200 is well ordered in the structure and the position of Cys247 could be also well predicted as it is close to the last ordered residue Pro244 in the structure. We have generated a model of the UL141 cysteine network to illustrate the possible intramolecular disulfide-bond interaction of UL141 (Supplementary Fig. S3).

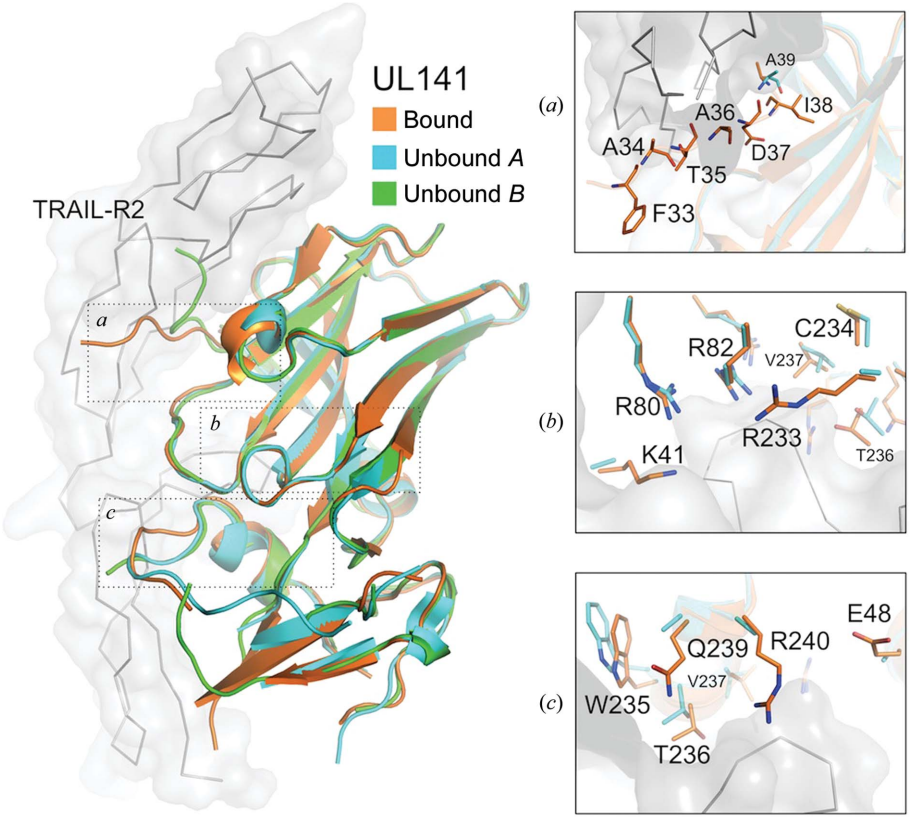

Figure 4
Superposition of bound and unbound UL141 structures. Superposition of C^{α} atoms of bound UL141 (chain *A* in orange, TRAIL-R2 chain *D* as a grey surface with ribbons; PDB entry 4ix9 chains *A* and *D*) and the unbound UL141 structure on the right (monomer *A* in cyan, monomer *B* in green; PDB entry 4jm0). The main structural differences are depicted in boxes (*a*, *b*, *c*) on the left. (*a*) The N-terminal loop shown by a stick model in orange (residues 33–39) is visible in the structure of UL141 (orange) bound to TRAIL-R2 (grey surface with ribbons) but is disordered in unliganded UL141 (cyan). (*b*, *c*) Two other examples of structural divergence between bound and unbound UL141 are indicated. (*b*) Interacting residues of the A′ and C β-strands are shifted or missing (Lys41, Arg233) when UL141 is unliganded. (*c*) The interacting α-helix H4 lacks side chains for residues Arg240, Gln239 and Glu48 of α-helix H1.

3.3. Structural comparison of bound and unbound UL141

Interpretation of the standard deviations between the bound and unbound structures is problematic owing to the low resolution and several missing regions in the C-terminal appendix domain of each monomer. To address this issue, we only calculated C^{α} deviations for well ordered regions. The Ig-like domain (residues 38–160, with 113 C^{α} atoms aligned) has an r.m.s.d. of 0.362 Å, and the C-terminal domain (residues 176–203, 210–215 and 231–243, with 36 C^{α} atoms aligned) has an r.m.s.d. of 0.582 Å. Superposition of the C^{α} atoms (Fig. 4) from UL141 bound to TRAIL-R2 (PDB entry 4ix9, orange) with those of unliganded UL141 (PDB entry 4jm0, monomer *A* in cyan, monomer *B* in green) indicates a high level of identity between the structures, with a total r.m.s.d. of 0.424 Å (155 C^{α} atoms aligned). However, we have

identified some mobile loops and residues that are not present in the unbound structure but are very well ordered in the structure of UL141 bound to TRAIL-R2. There are three notable examples of structural divergence between the structures. The first example is near the N-terminus (residues 30–38), within which unbound UL141 deviates significantly from UL141 bound to receptor. Not surprisingly, the solvent-exposed N-terminal loop is completely missing in monomer *B* and is also poorly ordered in monomer *A* (Fig. 4*a*). The reason for this difference could be attributed to the fact that the N-terminal loop of UL141 forms major contacts with the receptor (Nemčovičová *et al.*, 2013), while in the unbound state this loop is highly mobile and solvent-exposed. As illustrated in Fig. 4(*a*), this region contributes significantly to the binding of TRAIL-R2, which is in line with our previously reported mutagenesis study in this region of TRAIL-R2. Within the Ig-like domain there are two main regions of structural flexibility, one in the region of β-strands A′ and C (Fig. 4*b*) and the other in α-helix H4 (Fig. 4*c*). Both regions are involved in receptor binding. The side chains of residues Arg233, Lys41, Arg240, Gln239 and Glu48 as well as the abovementioned N-terminal loop (residues 30–37) show no density in the unbound structure but are very well ordered upon binding to TRAIL-R2. This suggests that these residues are flexible in solution and form important contacts with the receptor, marking them as good candidates for UL141 mutagenesis studies to assess binding to TRAIL-R2. Previously, mutational studies to characterize the binding regions of UL141 and TRAIL-R2 had only been performed on the receptor, not on UL141. UL141 mutants are the next logical step to compare their effect on binding to both TRAIL-R2 and its other cellular binding partner CD155.

3.4. Structural versatility of the UL141 Ig fold

While UL141 does not display any sequence homology to other proteins in the database, we previously performed a *DALI* search and identified significant structural conservation with other Ig-domain proteins, while no structurally related proteins were identified using the C-terminal domain of UL141 (Nemčovičová *et al.*, 2013). The second top hit from *DALI* was HCMV protein UL16, an immunoevasin that subverts NKG2D-mediated immune responses by retaining a select group of NKG2D ligands inside the cell (Müller *et al.*, 2010). As shown here (Fig. 5), UL16 aligns (yellow) with 85% of its structure to the Ig domain of UL141 (cyan; Z = 9.4, r.m.s.d. of 3.7 Å over 115 residues aligned). However, while UL16 binds to MHC-like molecules, UL141 has also evolved to target both TNFRs (TRAIL-R2; Fig. 5) and Ig-fold proteins (CD155; Fig. 5), illustrating the functional versatility of the Ig fold. In contrast to UL141, UL16 is heavily glycosylated (Müller *et al.*, 2010), which reduces the potential binding sites for cellular ligands. In order to visualize the native glycosylation of both UL16 and UL141, we show all N-linked glycans discovered by X-ray crystallography attached to asparagines as sphere representations (Fig. 5*a*). Six glycans shielding much

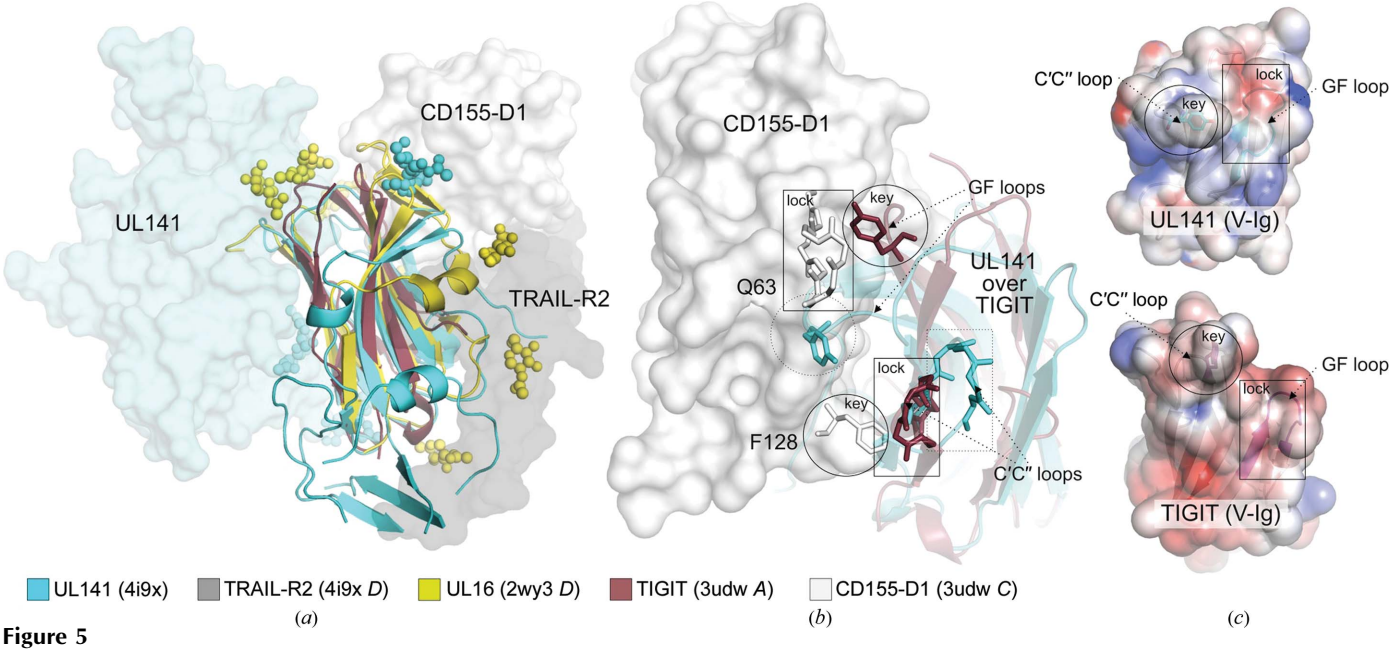

Figure 5
Superposition of HCMV UL141, HCMV UL16 and the human TIGIT–CD155-D1 complex and the electrostatic potential of UL141 and TIGIT. (*a*) Superposition of the Ig-like domains of the HCMV UL141–TRAIL-R2 complex (UL141 chain *A* is shown as a cyan cartoon, UL141 chain *B* as a cyan surface and TRAIL-R2 as a grey surface; PDB entry 4ix9 chains *A*, *B* and *D*), HCMV UL16 (yellow cartoon; PDB entry 2wy3 chain *D*; Müller *et al.*, 2010) and the TIGIT–CD155-D1 complex (TIGIT is shown as a raspberry red cartoon and CD155-D1 as a light grey surface; PDB entry 3udw chains *A* and *C*). (*b*) The positions of the residues involved in the 'lock-and-key' motif are indicated by sticks along the TIGIT–CD155-D1 and potential UL141–CD155-D1 interfaces. The C′C″ and GF loops important for CD155-D1 binding are depicted by arrows. Mutated residues in CD155 are labelled (Gln63 and Phe128). (*c*) Transparent surface over a cartoon representation of the dimerization interface of UL141 (top) and TIGIT (bottom). The molecular surfaces are coloured according to electrostatic potential. Positive potential is shown in blue and negative potential is shown in red; the potential is contoured from $-5kT\mathrm{e}^{-1}$ to $+5kT\mathrm{e}^{-1}$. The image was drawn using *PyMolX*11*Hybrid* (Schrödinger).

of the UL16 surface are shown as yellow spheres, while the three attached glycans of UL141 are shown in cyan. The overall glycan distribution on both proteins indicates that UL141 has a much more accessible surface for potential receptor binding than UL16, which only allows binding to a selected group of NKG2D ligands (such as MICB). This finding is in line with our previous prediction (Nemčovičová *et al.*, 2013) made by *ProMate* (Neuvirth *et al.*, 2004) for the potential CD155 binding site. Thus, UL141 has an obvious potential to engage more then one receptor at the time, while dimerization can also contribute to this function (Fig. 5*a*; the second monomer is shown as a cyan transparent surface).

We previously demonstrated that UL141 is capable of binding TRAIL-R2 and CD155 simultaneously (Nemčovičová *et al.*, 2013), indicating that UL141 has distinct binding sites for these two receptors. To further investigate the CD155 binding site in UL141, we searched the PDB for V-type Ig domains that also bind poliovirus receptor (CD155) or nectin-like molecules. Currently, there is only one crystal structure suitable for our analysis, which is the TIGIT (T-cell immunoreceptor with Ig and ITIM domains) molecule (Stengel *et al.*, 2012). Superposition of the N-terminal Ig-like domain of UL141 (residues 38–161; PDB entry 4ix9, chain *A*; Stengel *et al.*, 2012) with the Ig domain of TIGIT bound to CD155 (PDB entry 3udw, chain *A*) clearly indicates a potential binding site for CD155 on UL141 that is comprised of the GF and C′C″ loops (Figs. 5*b* and 5*c*).

3.5. Potential CD155 (domain D1) recognition by UL141

In the TIGIT–CD155-D1 complex structure, the TIGIT–CD155 interface is formed by interaction between the front β-sheet (A′GFCC′C″) of each molecule (Stengel *et al.*, 2012). Because the receptor and ligand share the same IgV fold, the interface displays approximate noncrystallographic twofold symmetry and is highly complementary in shape and charge. Interestingly, the interface in the TIGIT–CD155 complex uses the same structural elements as other IgV homodimers and heterodimers. Basically, the GF loop of each IgV domain contacts the C′C″ loop of its partner. As has been described previously (Stengel *et al.*, 2012), the conserved sequence motifs (A)X_6G in the C′C″ loop and (F/Y) in the GF loop define two signature lock-and-key interactions at symmetric corners of the interface that literally latch the two molecules together. According to the lock-and-key motifs described above, we were able to identify a potential binding site for CD155 in UL141. Similar to TIGIT, a potential concave 'lock' in the UL141 molecule could be formed by the LILDAVKG [(L)X_6G] motif (Supplementary Fig. S4; residues 93–100, orange) found in the C′C″ loop that creates a hydrophobic pocket (Fig. 5*c*). The potential convex 'key' could be formed by the aromatic residue Tyr148 in the FG loop of UL141 (Phe128 in CD155-D1) that could latch into the hydrophobic pocket (lock) on the opposing molecule (Fig. 5*b* and Supplementary Fig. S4). These lock-and-key motifs have been reported to be highly conserved in the IgV domain of nectins (but not Necls) and comprise the distinctive poliovirus receptor family motifs (Zhang *et al.*, 2008). However, sequence analysis reveals conservation of the (L/A)X_6G hydrophobic 'lock' and (Y/F) 'key' motifs in HCMV UL141 compared with the TIGIT and CD155 sequences (Supplementary Fig. S4, orange). Interestingly, mapping the positions of the mutants that affect the interaction of CD155 with poliovirus (Zhang *et al.*, 2008) onto the TIGIT–CD155 structure reveals that both poliovirus and TIGIT use the same surface and residues on CD155 for binding. This overlap of ligand- and virus-binding sites also has been reported for other virus receptors (Di Giovine *et al.*, 2011; Verdino *et al.*, 2010).

The proposed CD155 binding site (shown here by superposition of TIGIT on UL141) only slightly overlaps with one of the six binding patches defined for TRAIL-R2, in particular residue Tyr148. From alanine scanning, mutagenesis of the TRAIL-R2 residues L110A/L114A and F112A that target Tyr148 of UL141 reduced the binding to UL141 7.2-fold and 106-fold, respectively (Nemčovičová *et al.*, 2013). This suggests that this residue is important but does not abrogate TRAIL-R2 receptor binding to UL141 and could be potentially used for simultaneous binding to CD155.

3.6. Mutational analysis at the CD155–UL141 lock-and-key interface

To investigate the importance of the conserved CD155 motifs for formation of the lock-and-key complex, a couple of point mutants were created in CD155 (Fig. 5*b*). Each mutant protein was cloned as an Fc-fusion protein and expressed in 293T cells, and binding was determined using surface plasmon resonance (SPR). As reported previously (Nemčovičová *et al.*, 2013), wild-type CD155-Fc binds wild-type UL141 with high affinity (K_d = 1.9 n*M*). Reciprocally, when UL141-Fc was immobilized on the sensor chip and the binding of CD155 was assessed, the dissociation rate was roughly 2.6-fold slower, while the association rate was 1.7-fold faster, leading to a 4.3-fold lower equilibrium binding affinity (K_d = 8.5 n*M* *versus* 1.9 n*M*; Figs. 2*e* and 2*f*). The change in binding kinetics correlates with the notion that UL141 is a dimer in solution, while CD155 forms only monomers. We further reveal that the CD155 point mutant Q63R in the (L/A)X_6G motif and the 'key' region variant F128R in the (F/Y) motif abrogate binding to UL141 (Fig. 2*d*).

Taken together, we conclude that the lock-and-key interactions between CD155 and UL141 are critical and require the 'key' motif (F/Y) on CD155 for complex formation.

3.7. Dimerization interface of UL141

As we observed in the structure of UL141 bound to TRAIL-R2 (Nemčovičová *et al.*, 2013), structural analysis of unliganded UL141 revealed the same very tight interactions of well packed dimers in a head-to-tail fashion (Fig. 6 and Supplementary Fig. S5).

TNFR signalling is normally initiated following aggregation by trimeric TNF-family ligands (*e.g.* TRAIL); however, whether the noncanonical dimerization of TRAIL-R2 by UL141 can result in signalling, *e.g.* NF-κB activation for cell

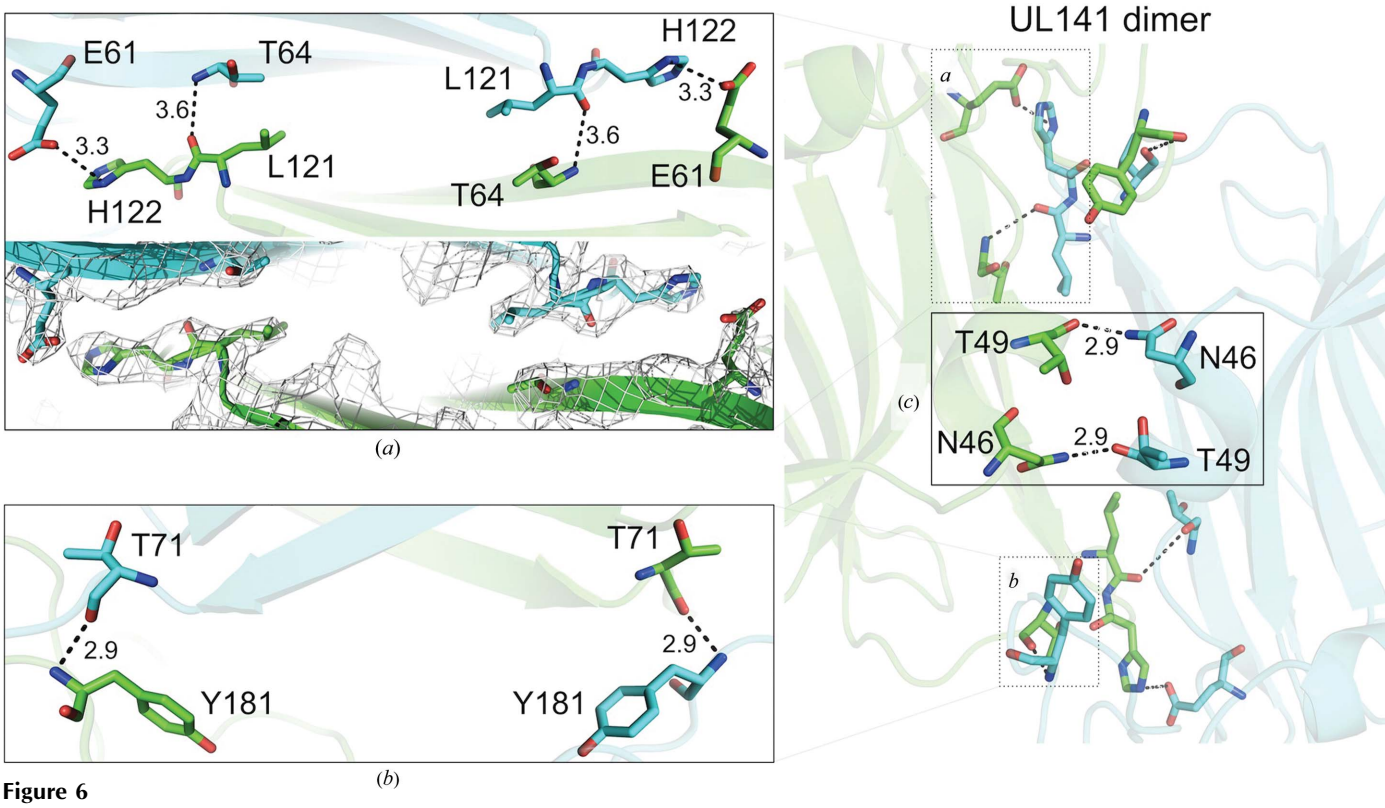

Figure 6
Dimerization interface of UL141. The positions of residues involved in dimerization (Asn46, Thr49, Glu61, Thr64, Thr71, Leu121, His122 and Tyr181) along the interface are indicated as sticks and the individual interactions are depicted in boxes (*a*, *b*, *c*). A stick model with a transparent cartoon in the background of the homodimer interface is shown in the 3.2 Å resolution solvent-flattened electron-density map contoured at 1σ (*a*, bottom) between monomer *A* (cyan) and *B* (green) of UL141. The main view centred on box (*c*) looks down the crystallographic twofold symmetry.

proliferation rather than death signalling, currently remains unclear.

The dimerization interface of UL141 is stabilized primarily through four hydrogen bonds and two salt bridges as well as several hydrophobic contacts, while burying a total surface area of 1423 Å^2. All these interactions occur symmetrically along the dimerization interface (Fig. 6). To better understand the character of the oligomerization of receptors on the membrane, we compared the dimerization and trimerization interfaces of HCMV UL141 and the endogenous ligand TRAIL, respectively (Supplementary Figs. S6*a* and S6*b*). The electrostatic potential revealed that both UL141 and TRAIL utilize their rather hydrophobic surfaces to interact with adjacent subunits and form fully functional oligomers.

To further investigate the dimeric properties of UL141, we generated a 'dimerization-deficient' mutant of UL141 (ddUL141) to test whether this disrupts the function and binding properties of UL141. The design of mutants disrupting the native oligomeric state of proteins is generally not straightforward, as many proteins, including TNF-family ligands, associate *via* hydrophobic surfaces to form functional oligomers (as shown in Supplementary Fig. S6*b* for TRAIL). Disruption of the native oligomeric state of a protein can often lead to aggregation, rather than resulting in a stable monomer. Therefore, we have randomly mutated various combinations of selected residues (Asn46, Glu61, Thr71 and Leu121) that form the UL141 dimer interface (Fig. 6 and Supplementary Fig. S5), and have assessed their potential to form a stable UL141 monomer by SEC. Using this approach, we were successful in producing a stable ddUL141 monomer by incorporating the mutations N46A, E61A and T71F (Fig. 2*b*). Based on the UL141–TRAIL-R2 structure, we hypothesize that this ddUL141 mutant should still be capable of interacting with TRAIL-R2. However, whether ddUL141 is capable of inhibiting cell-surface expression of TRAIL-R2, TRAIL-R1 or CD155 is currently unknown. In an effort to understand the mechanism by which HCMV UL141 functions on the surface of the cells, we sought to determine whether ddUL141 can bind to TRAIL-R2 and/or CD155 and form a stable complex. SPR analysis indicates that both UL141 and ddUL141 can bind to TRAIL-R2 (Fig. 2*a*) as well as to CD155 (Fig. 2*c*) with similar association rates as for the wild type, while dissociation is expectedly faster for ddUL141 owing to the loss of dimeric binding avidity that can be observed in the UL141 dimer (Figs. 2*a* and 2*c*). Whether dimerization of UL141 is indeed necessary to restrict cell-surface expression of the various host receptors is an important question that has yet to be addressed.

Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a Directorate of SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the US Department of Energy Office of Science by Stanford University. The SSRL Structural

Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393) and the National Center for Research Resources (P41RR001209). This publication is the result of the project implementation ITMS code 26240220071, supported by the Research and Development Operational Programme funded by the ERDF. In addition, we would like to thank the Institute of Neuroimmunology at the Slovak Academy of Sciences for allowing us to use the SPR instrumentation.

References

Ahn, K., Gruhler, A., Galocha, B., Jones, T. R., Wiertz, E. J., Ploegh, H. L., Peterson, P. A., Yang, Y. & Früh, K. (1997). *Immunity*, **6**, 613–621.

Alford, C. A., Stagno, S., Pass, R. F. & Britt, W. J. (1990). *Rev. Infect. Dis.* **12**, S745–S753.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.

Baumgarth, N., Choi, Y. S., Rothaeusler, K., Yang, Y. & Herzenberg, L. A. (2008). *Curr. Top. Microbiol. Immunol.* **319**, 41–61.

Beck, S. & Barrell, B. G. (1988). *Nature (London)*, **331**, 269–272.

Beer, T. A. de, Berka, K., Thornton, J. M. & Laskowski, R. A. (2014). *Nucleic Acids Res.* **42**, D292–D296.

Benedict, C. A., Banks, T. A. & Ware, C. F. (2003). *Curr. Opin. Immunol.* **15**, 59–65.

Bottino, C., Castriconi, R., Pende, D., Rivera, P., Nanni, M., Carnemolla, B., Cantoni, C., Grassi, J., Marcenaro, S., Reymond, N., Vitale, M., Moretta, L., Lopez, M. & Moretta, A. (2003). *J. Exp. Med.* **198**, 557–567.

Cha, T.-A., Kao, K., Zhao, J., Fast, P. E., Mendelman, P. M. & Arvin, A. (2000). *J. Clin. Microbiol.* **38**, 839–845.

Cha, T.-A., Tom, E., Kemble, G. W., Duke, G. M., Mocarski, E. S. & Spaete, R. R. (1996). *J. Virol.* **70**, 78–83.

Chalupny, N. J., Rein-Weston, A., Dosch, S. & Cosman, D. (2006). *Biochem. Biophys. Res. Commun.* **346**, 175–181.

Chapman, T. L., Heikeman, A. P. & Bjorkman, P. J. (1999). *Immunity*, **11**, 603–613.

Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., Horsnell, T., Hutchison, C. A. III, Kouzarides, T., Martignetti, J. A. & Barrell, B. G. (1990). *Curr. Top. Microbiol. Immunol.* **154**, 125–169.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Cohen, S. X., Ben Jelloul, M., Long, F., Vagin, A., Knipscheer, P., Lebbink, J., Sixma, T. K., Lamzin, V. S., Murshudov, G. N. & Perrakis, A. (2008). *Acta Cryst.* D**64**, 49–60.

Cosman, D., Müllberg, J., Sutherland, C. L., Chin, W., Armitage, R., Fanslow, W., Kubin, M. & Chalupny, N. J. (2001). *Immunity*, **14**, 123–133.

Davison, A. J., Dolan, A., Akter, P., Addison, C., Dargan, D. J., Alcendor, D. J., McGeoch, D. J. & Hayward, G. S. (2003). *J. Gen. Virol.* **84**, 17–28.

Di Giovine, P., Settembre, E. C., Bhargava, A. K., Luftig, M. A., Lou, H., Cohen, G. H., Eisenberg, R. J., Krummenacher, C. & Carfi, A. (2011). *PLoS Pathog.* **7**, e1002277.

Dolan, A., Cunningham, C., Hector, R. D., Hassan-Walker, A. F., Lee, L., Addison, C., Dargan, D. J., McGeoch, D. J., Gatherer, D., Emery, V. C., Griffiths, P. D., Sinzger, C., McSharry, B. P., Wilkinson, G. W. & Davison, A. J. (2004). *J. Gen. Virol.* **85**, 1301–1312.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Fuchs, A., Cella, M., Giurisato, E., Shaw, A. S. & Colonna, M. (2004). *J. Immunol.* **172**, 3994–3998.

Furman, M. H., Dey, N., Tortorella, D. & Ploegh, H. L. (2002). *J. Virol.* **76**, 11753–11756.

Grosse-Kunstleve, R. W. & Adams, P. D. (2002). *Acta Cryst.* A**58**, 60–65.

Ho, M. (1991). *Transplant. Proc.* **23**, 2–7.

Ho, K. L., Gottlieb, C. & Zarbo, R. J. (1991). *Clin. Neuropathol.* **10**, 127–133.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Lanier, L. L. (2008). *Nature Rev. Immunol.* **8**, 259–268.

Laskowski, R. A. (2009). *Nucleic Acids Res.* **37**, D355–D359.

Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). *J. Mol. Biol.* **231**, 1049–1067.

Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart, D. S. (2004). *Nucleic Acids Res.* **32**, W590–W594.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Müller, S., Zocher, G., Steinle, A. & Stehle, T. (2010). *PLoS Pathog.* **6**, e1000723.

Murphy, E., Yu, D., Grimwood, J., Schmutz, J., Dickson, M., Jarvis, M. A., Hahn, G., Nelson, J. A., Myers, R. M. & Shenk, T. E. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 14976–14981.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nemčovičová, I., Benedict, C. A. & Zajonc, D. M. (2013). *PLoS Pathog.* **9**, e1003224.

Neuvirth, H., Raz, R. & Schreiber, G. (2004). *J. Mol. Biol.* **338**, 181–199.

Painter, J. & Merritt, E. A. (2006). *Acta Cryst.* D**62**, 439–450.

Pass, R. F. (2005). *Herpes*, **12**, 50–55.

Prod'homme, V., Griffin, C., Aicheler, R. J., Wang, E. C. Y., McSharry, B. P., Rickards, C. R., Stanton, R. J., Borysiewicz, L. K., López-Botet, M., Wilkinson, G. W. & Tomasec, P. (2007). *J. Immunol.* **178**, 4473–4481.

Prod'homme, V., Sugrue, D. M., Stanton, R. J., Nomoto, A., Davies, J., Rickards, C. R., Cochrane, D., Moore, M., Wilkinson, G. W. & Tomasec, P. (2010). *J. Gen. Virol.* **91**, 2034–2039.

Roy, C. R. & Mocarski, E. S. (2007). *Nature Immunol.* **8**, 1179–1187.

Sinzger, C., Grefte, A., Plachter, B., Gouw, A. S. H., The, T. H. & Jahn, G. (1995). *J. Gen. Virol.* **76**, 741–750.

Smith, W. *et al.* (2013). *Cell Host Microbe*, **13**, 324–335.

Stengel, K. F., Harden-Bowles, K., Yu, X., Rouge, L., Yin, J., Comps-Agrar, L., Wiesmann, C., Bazan, J. F., Eaton, D. L. & Grogan, J. L. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 5399–5404.

Stern-Ginossar, N. *et al.* (2007). *Science*, **317**, 376–381.

Stern-Ginossar, N., Weisburd, B., Michalski, A., Le, V. T. K., Hein, M. Y., Huang, S.-X., Ma, M., Shen, B., Qian, S.-B., Hengel, H., Mann, M., Ingolia, N. T. & Weissman, J. S. (2012). *Science*, **338**, 1088–1093.

Tomasec, P., Braud, V. M., Rickards, C., Powell, M. B., McSharry, B. P., Gadola, S., Cerundolo, V., Borysiewicz, L. K., McMichael, A. J. & Wilkinson, G. W. G. (2000). *Science*, **287**, 1031–1033.

Tomasec, P., Wang, E. C. Y., Davison, A. J., Vojtesek, B., Armstrong, M., Griffin, C., McSharry, B. P., Morris, R. J., Llewellyn-Lacey, S., Rickards, C., Nomoto, A., Sinzger, C. & Wilkinson, G. W. G. (2005). *Nature Immunol.* **6**, 181–188.

Ulbrecht, M., Martinozzi, S., Grzeschik, M., Hengel, H., Ellwart, J. W., Pla, M. & Weiss, E. H. (2000). *J. Immunol.* **164**, 5019–5022.

Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.

Verdino, P., Witherden, D. A., Havran, W. L. & Wilson, I. A. (2010). *Science*, **329**, 1210–1214.

Wang, Z., Turner, R., Baker, B. M. & Biddison, W. E. (2002). *J. Immunol.* **169**, 3146–3154.
Ware, C. F. & Sedý, J. R. (2011). *Curr. Opin. Immunol.* **23**, 627–631.
Wills, M. R., Ashiru, O., Reeves, M. B., Okecha, G., Trowsdale, J., Tomasec, P., Wilkinson, G. W., Sinclair, J. & Sissons, J. G. (2005). *J. Immunol.* **175**, 7457–7465.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yang, Z. & Bjorkman, P. J. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 10095–10100.
Zhang, K. Y. J., Cowtan, K. & Main, P. (1997). *Methods Enzymol.* **277**, 53–64.
Zhang, P., Mueller, S., Morais, M. C., Bator, C. M., Bowman, V. D., Hafenstein, S., Wimmer, E. & Rossmann, M. G. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 18284–18289.