REGULATION OF THE CELL SURFACE EXPRESSION OF CLASSICAL AND NON-CLASSICAL MHC PROTEINS BY THE HUMAN CYTOMEGALOVIRUS UL40 AND RHESUS CYTOMEGALOVIRUS RH67 PROTEINS

Simon Brackenridge[#], Persephone Borrow, and Andrew McMichael

Nuffield Department of Clinical Medicine, University of Oxford, Oxford, United Kingdom

[#]Address correspondence to Simon Brackenridge, simon.brackenridge@ndm.ox.ac.uk

Running Head: HCMV UL40 and RhCMV Rh67 regulate MHC expression level

ABSTRACT: 250 words

IMPORTANCE: 148 words

INTRODUCTION:, METHODS & DISCUSSION: 5888 words

ABSTRACT

The signal sequences of the human cytomegalovirus (CMV) UL40 protein, and its rhesus CMV counterpart, Rh67, contain a peptide (VMAPRT[L/V][F/I/L/V]L, VL9) that can be presented by Major Histocompatibility Complex (MHC) antigen E. The CMV VL9 peptides replace VL9 peptides derived from classical MHC (Ia) signal sequences, which are lost when CMV disrupts MHC Ia expression, as well as antigen processing and presentation. This allows infected cells to maintain surface expression of MHC-E and escape killing by NK cells expressing the inhibitory NKG2A/CD94 receptor. We demonstrate that processing of the Rh67 VL9 peptide mirrors that of UL40, despite the lack of sequence conservation elsewhere in the two proteins. As previously shown for UL40, up-regulation of MHC-E expression by Rh67 only requires its signal sequence, with sequences upstream of VL9 critical for conferring independence from TAP, the Transporter Associated with Antigen Processing. Additionally, we show that processing of both VL9 peptides depends on cleavage of the signal sequences by the host protease Signal Peptide Peptidase. Notably, our results also reveal that the mature UL40 and Rh67 proteins contribute to CMV immune evasion by down-regulating surface expression of MHC Ia. Unexpectedly, while the Rh67 VL9 peptide is resistant to the effects of the Rh67 protein, the UL40 protein is able to counteract up-regulation of MHC-E expression mediated by its own VL9 peptide. This suggests differences in the mechanisms by which the two VL9 peptides up-regulate MHC-E, which may have implications for translating a RhCMV-vectored SIV vaccine to HIV-1 using HCMV as a vector.

34 # IMPORTANCE

35 Cell surface MHC-E expression, which requires a peptide (VMAPRT[L/V][F/I/L/V]L, "VL9")
36 from the signal sequences of other MHC class I proteins, prevents cells from being killed by
37 CD94/NKG2A-expressing Natural Killer cells. In cells infected with human CMV, the
38 endogenous VL9 peptide is replaced by one from the signal sequence of the HCMV UL40
39 protein. We show that processing of the VL9 peptide of Rh67, the rhesus CMV equivalent of
40 UL40, mirrors that of UL40, despite a lack of sequence homology between the two proteins.
41 Of note, we also show that the mature UL40 and Rh67 proteins, which have no previously
42 described function, contribute to CMV immune evasion by reducing classical MHC class I
43 surface expression. Importantly, the mature UL40 protein, but not the mature Rh67
44 protein, can decrease the up-regulation of MHC-E mediated by its signal sequence VL9
45 peptide, which may have implications for HCMV as a vaccine vector.

46 # INTRODUCTION

47 Cell surface expression of the Major Histocompatibility (MHC) E protein, known as HLA-E in
48 humans, depends on a nonamer peptide (VMAPRT[L/V][F/I/L/V]L, VL9) present in the signal
49 sequences of many MHC class Ia (MHC Ia) allotypes (1, 2). Presentation of this peptide is
50 dependent on cleavage by both the endoplasmic reticulum (ER) membrane resident
51 proteases Signal Peptidase (SP) and Signal Peptide Peptidase (SPP) (3, 4) and the cytosolic
52 proteasome (5), import into the ER lumen by the transporter associated with antigen
53 presentation (TAP), as well as tapasin and the peptide loading complex (PLC) (6).

54 The presence of HLA-E at the cell surface is monitored by NK cells via CD94/NKG2 receptor
55 complexes (7, 8) as a proxy for MHC class I expression. Loss of surface HLA-E results in the
56 loss of signalling through the inhibitory NKG2A/CD94 complex, allowing NK cells to kill HLA-
57 E-deficient cells (9). Cell surface expression of HLA-E is, therefore, often maintained by
58 virally infected cells, in contrast to that of classical MHC class I proteins. For example,
59 evasion of $CD8^+$ T cell responses by human CMV (HCMV) involves not only direct targeting of
60 the MHC Ia proteins themselves, including retention in the ER by US10 (10) and targeting for
61 degradation by US2 and US11 (11, 12), but also targeting of accessory factors. US6 blocks
62 peptide transport by TAP (13), while US3 inhibits tapasin (14). Such comprehensive
63 disruption of both MHC Ia expression and components of the PLC limits the availability of
64 the VL9 peptide for HLA-E, resulting in NK cell-mediated killing of HCMV infected cells.
65 However, the signal sequence of the HCMV UL40 protein contains a VL9 peptide, which can
66 be presented by HLA-E in a TAP-independent manner (15, 16), allowing HCMV-infected cells
67 to maintain surface expression of HLA-E and avoid detection by NK cells.

68 Immune evasion by rhesus CMV (RhCMV) appears to be highly conserved, and equivalents
69 of the HCMV US2, US3, US6, and US11 proteins have been identified that function by
70 conserved mechanisms: Rh182, Rh184, Rh186, and Rh188, respectively (17). RhCMV also
71 appears to evade NK cells in the same way as HCMV, via the Rh67 protein. Despite
72 otherwise appearing unrelated to UL40, Rh67 also contains the VL9 peptide in its signal
73 sequence (18), and we previously showed that the Rh67 VL9 peptide can stabilise HLA-E at
74 the cell surface (19). More recently, it has recently been demonstrated that presentation of
75 the Rh67 VL9 peptide does not require TAP (20), confirming that Rh67 is indeed the RhCMV
76 equivalent of UL40.

77 A RhCMV-vectored vaccine containing the SIV Gag, Rev/Tat/Nef, Env and Pol sequences
78 elicits an immune response that enables ~55% of vaccinated rhesus macaques to clear SIV
79 infection early after virus challenge (21, 22). Unusually, the CD8+ T cell responses induced
80 by this vaccine are not restricted by classical MHC Ia proteins. Instead, approximately two-
81 thirds of the $CD8^+$ T cells are restricted by MHC class II, with the rest restricted by Mamu-E,
82 the rhesus macaque homologue of HLA-E (19). A full understanding of exactly how
83 protection is mediated by this vaccine remains to be elucidated, but it has recently been
84 shown to depend on the Mamu-E-restricted responses rather than the MHC class II-
85 restricted responses (23, 24).

86 Successful development of a HCMV-vectored HIV-1 vaccine will likely depend on a detailed
87 understanding of the similarities and differences in the immune evasion strategies
88 employed by RhCMV and HCMV. Indeed, such an understanding may even allow the use of
89 HCMV to be avoided, with the required CMV genes being engineered into a more
90 conventional vaccine vector. To this end, we have undertaken a detailed comparison of

91 how the VL9 peptides from UL40 and Rh67 are processed for presentation by HLA-E. Our
92 results reveal that the similarities between UL40 and Rh67 extend beyond their
93 independence from TAP. Processing of both the UL40 and Rh67 VL9 peptides is dependent
94 on cleavage by SPP, and the up-regulation of HLA-E mediated by Rh67 requires only the
95 signal sequence, as was reported previously for UL40 (16). Furthermore, we show that the
96 mature Rh67 and UL40 proteins also contribute to the evasion of the host immune response
97 by decreasing surface expression of MHC Ia. Unexpectedly, the mature UL40 and Rh67
98 proteins are also capable of counteracting the up-regulation of HLA-E mediated by the UL40
99 signal sequence, but not the Rh67 signal sequence. This suggests that the TAP-
100 independence of the UL40 and Rh67 VL9 peptides may be achieved by distinct mechanisms,
101 and that there may be differences in the level of MHC-E expression required to evade NK
102 cells in humans and macaques.

103 METHODS

104 *Expression Plasmids and cloning*

105 The single chain dimer (SCD) of HLA-E*01:01 and β2-microglobulin was as previously
106 described (19), and the HLA-E*01:03 SCD was made by an identical cloning strategy. The
107 coding sequences of HLA-A*02:01 (with its own signal sequence, or the signal sequences of
108 UL40 or Rh67), KIR2DS2*001 (used as the irrelevant control), UL40 (HCMV strain Toledo),
109 and Rh67 (RhCMV strain 68-1) were purchased as synthetic genes from Eurofins with
110 appropriate restriction sites (HindIII and BamHI) and inserted into pEGFP-N1 (Clontech).
111 Constructs to express peptides in the cytosol were created by inserting annealed
112 oligonucleotides comprising a Kozak sequence, a start codon, the coding sequence of the
113 peptide and a stop codon, flanked by the overhangs for HindIII and AgeI, into pEGFP-N1 cut

114 with the same enzymes. For the constructs expressing peptides with a signal sequence, a
115 modified version of the signal sequence of HLA-E (with the final codon for alanine [GCG]
116 replaced by one for proline [CCA] to create an MscI site [TGGCCA]) was first inserted into
117 pEGFP-N1 using HindIII and AgeI. Annealed oligos encoding alanine followed by the desired
118 peptide and a stop codon, and with the overhang for AgeI at the 3' end, were inserted after
119 the signal sequence using the MscI and AgeI restriction sites. SPP expression constructs
120 were made by inserting cDNAs amplified from 293T cells (see below) into pcDNA3.1neo (Life
121 Technologies) using the HindIII and NotI sites included in the PCR primers. The catalytically-
122 inactive SPP D219A mutant was created using overlap-extension PCR (25) with KOD Hot
123 Start Polymerase (Merck) and appropriate primers. PCR products were purified using
124 QIAquick spin columns (QIAGEN). The sequences of the oligonucleotides used are shown in
125 Table S1. The HCV Core-E1 expression plasmid was a kind gift of Andrea Magri (University of
126 Oxford). All plasmids were prepared using QIAprep mini spin columns (QIAGEN), and
127 sequences were verified by Sanger sequencing using an ABI3770.

128 *Cell culture and transfection*

129 HEK 293T cells were maintained between 10% and 90% confluency at 37°C/5% CO2 in
130 DMEM (Life Technologies) supplemented with 10% foetal bovine serum (Sigma), and
131 penicillin/streptomycin (50 units/ml and 50 µg/ml, respectively; Life Technologies).
132 Transfections were carried out in 6 well plates with cells at 50–70% confluency using
133 GeneJuice (Merck) with 1 µg of total plasmid DNA per well, as per the manufacturer's
134 instructions.

135 *Gene inactivation by CRISPR/Cas9*

136 Guide RNAs for TAP1, TAP2, and HM13 (Table S2) were selected from a previously published
137 set (26) and inserted into pspgRNA (Addgene plasmid # 47108, a gift from Charles Gersbach)

138 (27). HEK 293T cells were co-transfected with equal amounts of the pspgRNA plasmids and
139 pCas9_GFP (Addgene plasmid #44719, a gift from Kiran Musunuru) (28). EGFP-positive
140 single cells were sorted 48 hours post transfection, and genomic DNA isolated using
141 QuickExtract (Lucigen). The gRNA binding sites were amplified using KOD Hot Start
142 Polymerase (Merck) and the primers shown in Table S3, with the following cycling
143 conditions: 1 minute at 96°C; 5 cycles of 25 seconds at 96°C, 45 seconds at 70°C, 45 seconds
144 at 72°C; 21 cycles of 25 seconds at 96°C, 50 seconds at 65°C, 45 seconds at 72°C; 4 cycles of
145 25 seconds at 96°C, 60 seconds at 55°C, 120 seconds at 72°C. PCR products were resolved
146 on a 1.5% agarose gel and visualised with SYBRSafe (Life Technologies), purified using
147 QIAquick spin columns (QIAGEN) and sequenced directly with both amplification primers.
148 The mutations in traduced by Cas9 were deconvoluted manually or using TIDE
149 (https://tide.deskgen.com/) (29). Deconvolutions were confirmed by cloning the PCR
150 products using a ZERO BLUNT PCR cloning kit (Life Technologies) and sequencing individual
151 clones.

152 *Protein extraction and Western blotting*

153 Cell lysates were prepared by incubating cells in RIPA buffer (150mM NaCl, 5mM EDTA,
154 50mM TRIS-Cl pH 8.0, 1% IGEPAL CA-630, 0.5% Sodium Deoxycholate, 0.1% SDS) on ice for
155 20 minutes followed by centrifugation at 13000 rpm (4 °C) for 10 minutes to remove
156 organelle debris. Equal volumes of lysate were mixed with 2x LDS Loading buffer (Life
157 Technologies) and run on 10% Bis-Tris NuPage gels (Life Technologies) in 1x MOPS running
158 buffer, then transferred to PVDF membrane using 2x NOVEX Transfer Buffer (Life
159 Technologies) and a Trans-Blot SD semi-dry transfer cell (Bio-Rad). Membrane blocking and
160 antibody incubations were done at room temperature with constant rotation, in 5% non-fat
161 milk prepared from powder in PBS/0.1% Tween-20, and were washed three times (5

162 minutes each) after each antibody incubation with PBS/0.1% Tween-20. Rabbit anti-SPP
163 (ab190253, Abcam) or mouse anti-Hepatitis C virus Core protein (clone C7-50, Life
164 Technologies) were visualised using IRDye 800CW Donkey anti-Rabbit IgG or Goat anti-
165 Mouse IgG (LiCor). The control antibody was DyLight 680 β-Tubulin Loading Control
166 Monoclonal Antibody (Life Technologies). Membranes were imaged using a LICOR Odyssey
167 X.

168 *HM13 Reverse-transcription PCR*

169 1µg of total cellular RNA prepared using RNeasy mini columns (QIAGEN) was reverse
170 transcribed using EvoScript Universal cDNA Master (ROCHE). PCR amplification using the
171 primers shown in Table S1, and electrophoresis, purification and cloning of the PCR products
172 was as before.

173 *Staining & flow cytometry*

174 HEK293T cells were stained 24 hours post transfection in 100µl of Dulbecco's modified PBS
175 (DPBS, Sigma) at 4°C for 15 minutes with antibody 3D12 (BioLegend) at 5 ng/µl, washed
176 twice with DPBS, stained with secondary antibody (allophycocyanin-crosslinked Goat-Anti-
177 Mouse (H+L) F(ab')2 fragment, Life Technologies) at 130 pg/µl in 100µl of DPBS for 15
178 minutes at 4°C, washed as before, and fixed in 100µl of Cytofix (BD Biosciences). Stained
179 cells were acquired using a CyAn ADP Analyser (Beckman Coulter), and analysed using
180 FlowJo 10 (BD).

181 # RESULTS

182 *An experimental system to dissect up-regulation of HLA-E by UL40 and Rh67*

183 Endogenous cell surface expression of HLA-E is very low, so we adopted the strategy we
184 used previously to show that Rh67 could increase surface expression of HLA-E (19). 293T

185 cells were transiently transfected with a plasmid expressing a single chain dimer (SCD) of
186 HLA-E*01:01 and β2-microglobulin (Figure S1A) and expression was detected using the HLA-
187 E-specific monoclonal antibody 3D12 (the gating strategy is shown in Figure S1B).
188 Consistent with previous reports regarding the relative expression of the two common HLA-
189 E alleles (30), this HLA-E*01:01 SCD expressed at lower levels than the HLA-E*01:03 SCD
190 (Figure S1C), confirming that covalently linking β2-microglobulin to HLA-E does not
191 significantly alter expression of HLA-E. Co-transfection of plasmids expressing both UL40
192 and Rh67 (protein sequences shown in Figure 1A) with the HLA-E*01:01 SCD increased cell
193 surface expression of HLA-E (Figure 1B). A more modest, but reproducible, increase in
194 surface expression was also seen when a plasmid expressing the HLA-A*02:01 protein is co-
195 transfected. The up-regulation of HLA-E by Rh67 was always slightly greater than that
196 mediated by UL40, a difference that may simply reflect the fact that the Rh67 signal
197 sequence appears to direct translation to the ER better than that of UL40 (Figure S2). In
198 contrast, the lower up-regulation mediated by HLA-A*02 presumably stems from its VL9
199 peptide being processed less efficiently than the VL9 peptides of UL40 and Rh67 as the HLA-
200 A*02 signal sequence is equivalent to the UL40 signal sequence in directing translation to
201 the ER.

202 The UL40 mediated up-regulation of HLA-E expression has long been known not to require
203 TAP (15, 16), and it has recently been shown that this is also true for the increase in MHC-E
204 expression induced by Rh67 (20). To confirm that both UL40 and Rh67 were up-regulating
205 HLA-E expression by a TAP-independent pathway in our 293T cells, we used CRISPR/Cas9 to
206 separately inactivate the *TAP1* or *TAP2* genes (see Figure S3 and Tables S4 and S5 for
207 details). For each, three single cell clones were established that had reduced MHC Ia
208 expression consistent with the inactivation of TAP. Of these, clones TAP1-1 and TAP2-4

209 (genetic lesions shown in Figure S3E and F) were confirmed TAP deficient (Figure S3G, H and
210 I). For both the TAP1-1 and TAP2-4 clones, we observed an almost total abrogation of the
211 HLA-A*02-mediated up-regulation of HLA-E expression, while UL40 and Rh67 were both still
212 able to mediate significant increases in the surface expression of HLA-E (Figure 1C). This
213 confirms that delivery of the Rh67 peptide to HLA-E does not require transit from the
214 cytosol to the ER lumen via TAP in 293T cells, allowing these cells to be used to dissect the
215 sequences required for the TAP-independent increase in HLA-E expression mediated by
216 Rh67.

217 *The N-terminal portion of the Rh67 signal sequence is critical for TAP-independent up-*
218 *regulation of HLA-E*

219 Up-regulation of HLA-E by UL40 is known to only require the UL40 signal sequence, and the
220 13 amino acids between the start codon and the UL40 VL9 peptide are both necessary for
221 TAP-independent up-regulation of HLA-E by the UL40 VL9 peptide, and sufficient to confer
222 TAP independence on the VL9 peptide present in the HLA-A*02 signal sequence (16) (Figure
223 S4). As expected, replacing the Rh67 signal sequence with that of HLA-E (which does not
224 contain the VL9 peptide) abolished up-regulation of HLA-E (Figure 2A, ESS-Rh67), whereas
225 replacing the mature Rh67 coding sequence with that of HLA-A*02 did not (Figure 2B, RSS-
226 A2). These results confirm that the signal sequence of Rh67 is sufficient to mediate the up-
227 regulation of HLA-E surface expression, and that the mature protein is not required.

228 The Rh67 signal sequence has only has 6 amino acids between its start codon and its VL9
229 peptide, and this sequence is less polar than the equivalent 13 amino acids in UL40. Despite
230 this, replacing the 6 intervening amino acids in the Rh67 signal sequence with a single
231 arginine (to preserve the expected charge distribution, as was done previously with UL40)
232 abolished up-regulation of HLA-E by Rh67 in both the *TAP1* and *TAP2* knockout cells

233 (Rh67ΔN, Figure 2C), without significantly affecting the functioning of the Rh67 signal
234 sequence (Figure S2). Conversely, replacing the two amino acids upstream of the HLA-A*02
235 VL9 peptide with the first 7 amino acids from the N-terminus of the Rh67 signal sequence
236 allowed TAP-independent up-regulation of HLA-E expression by the HLA-A*02 VL9 peptide
237 (RVA-A2; Figure 2D). Thus, despite differing in both sequence composition and length from
238 the corresponding sequence in UL40, the N-terminus of the Rh67 signal sequence is both
239 necessary for the TAP-independence of the Rh67 VL9 peptide, and sufficient to confer TAP-
240 independence on the HLA-A*02 VL9 peptide.

241 *Signal peptide peptidase is required for the processing of the UL40 and Rh67 VL9*
242 *peptides*

243 Processing of the VL9 peptide from the MHC Ia signal sequences is known to require
244 cleavage of the signal sequence downstream of the VL9 peptide by SPP (3). Although it was
245 initially reported that the UL40 signal sequence was not a substrate for SPP (31), a later
246 report suggested that it could be inefficiently processed by SPP or a SPP-like protease (16).
247 However, it is not clear if cleavage by SPP is required to generate the UL40 VL9 peptide.
248 Indeed, the UL40 signal sequence may have evolved to be a poor substrate if cleavage by
249 SPP prevents proper processing of the UL40 VL9 peptide. Therefore, to address whether
250 SPP is required for processing of the VL9 peptides in the UL40 and Rh67 signal sequences,
251 we used CRISPR/Cas9 to knock out *HM13*, the gene that encodes SPP.

252 The *HM13* gRNA used targets the end of exon 4 (with the Protospacer Adjacent Motif
253 overlapping the 5' splice site), which encodes the longest cytoplasmic loop of SPP (Figure
254 S5A). Four single cell clones were selected that had no SPP protein detectable by western
255 blot (Figure S5B). Three of these clones had lesions in the *HM13* gene that would be
256 predicted to have significant effects on expression of functional protein (Figure S5C; Table

257 S6). The exception (clone HM13-2) had a 10 bp deletion in one allele (resulting in
258 premature termination of translation in exon 5), and the in-frame deletion of the final 9 bp
259 of the exon in the other. Although this clone could still express SPP with a 3 amino acid
260 deletion, the altered sequence at the end of the exon is predicted to have a significant
261 effect on recognition of the adjacent 5' splice site (Table S7), likely resulting in inefficient or
262 aberrant splicing that may well affect expression of the SPP protein.

263 When the *HM13* knock out clones were screened for their ability to support HLA-E up-
264 regulation by UL40 and Rh67, two distinct phenotypes were observed (Figure 3A). None of
265 the clones showed up-regulation of HLA-E expression by UL40, while two (HM13-1 and
266 HM13-2) supported up-regulation by Rh67. Clones HM13-3 and HM13-4 were also unable
267 to support maturation of the HCV Core-E1 precursor protein, which is cleaved by SPP (32),
268 with only the immature iCore protein being detected (Figure 3B). In contrast, clone HM13-1
269 was indistinguishable from the parental cells, with only the mature Core (mCore) protein
270 observed. Clone HM13-2 had an intermediate phenotype, with both mCore and iCore being
271 detected. We conclude that clones HM13-1 and HM13-2 both retain some form of SPP
272 activity, while clones HM13-3 and HM13-4 are truly SPP deficient.

273 The failure of UL40 and Rh67 to up-regulate expression of HLA-E in clones HM13-3 and
274 HM13-4 suggests that processing of both the UL40 and Rh67 VL9 peptides is dependent on
275 SPP. To confirm this, we attempted to rescue the up-regulation of HLA-E by transiently
276 over-expressing SPP. Expression of wild type SPP in clone HM13-3 rescued up-regulation of
277 HLA-E by both UL40 and Rh67 (Figure 3C, top row), and also cleavage of the HCV Core-E1
278 precursor (Figure 3D). The up-regulation of HLA-E by UL40 was more sensitive to the
279 amount of SPP expression plasmid used in these transfections than that of Rh67 (Figure 3C,

280 rows 2 and 3), consistent with the UL40 signal sequence being processed inefficiently by
281 SPP. A catalytically inactive mutant of SPP (with the key aspartic acid residue at position 219
282 mutated to alanine) did not restore processing of the HCV Core-E1 precursor (Figure 3D), or
283 up-regulation of HLA-E by UL40 and Rh67 (Figure 3E).

284 These results confirm that cleavage by SPP is essential for generating the VL9 peptides from
285 both the UL40 and Rh67 signal sequences. Up-regulation of HLA-E expression by Rh67 in
286 clones HM13-1 and HM13-2 could result from these cells retaining low levels of SPP activity
287 (insufficient to support processing of UL40), or expressing SPP with altered substrate
288 specificity as a result of the CRISPR/Cas9 lesions. It proved possible to amplify SPP cDNA
289 from clone HM13-1, and a number of novel isoforms were identified by sequencing (Figure
290 S6A; Table S8). The predominant isoform (HM13-1 isoform D) was almost identical to SPP
291 isoform 1 and encoded an aberrant variant of SPP with the sequence of exon 4 (which
292 encodes the longest cytoplasmic loop) replaced by a cryptic exon from within intron 3
293 (Figure S6B). However, over-expression of this isoform in clone HM13-3 rescued up-
294 regulation of HLA-E by both UL40 and Rh67, confirming that this novel isoform could
295 recognise both signal sequences (Figure S6C). Therefore, the impairment to UL40 mediated
296 up-regulation in clone HM13-1 (and HM13-2) most likely stems from the low level of
297 expression of this novel isoform, sufficient to allow processing of the VL9 peptide of Rh67,
298 but not that of UL40.

299 *Improving cleavage of the UL40 signal sequence by SPP.*

300 Consistent with the UL40 signal sequence being a poor substrate for SPP, over-expression of
301 wild type SPP in the parental 293T cells resulted in a modest increase in up-regulation of
302 HLA-E by only UL40 (Figure 4A), suggesting that the up-regulation of surface expression by
303 UL40 may be limited by inefficient SPP cleavage. To confirm this, we tested the effects of a

304 previously described double mutant of the UL40 signal sequence (R31T and R33T (31) –
305 Figure 4B) that was suggested to convert the UL40 signal sequence into an efficient SPP
306 substrate. This mutation had little effect on the functioning of the signal sequence (Figure
307 S2), so presumably has a direct effect on cleavage by SPP, rather than improving cleavage
308 simply by targeting the signal sequence to the ER more efficiently. Consistent with an
309 improvement in processing by SPP, this mutant was capable of upregulating HLA-E
310 expression in clone HM13-1 (Figure 4C, left-hand panel). In the parental 293T cells,
311 expression of the mutant UL40 increased up-regulation of HLA-E more than the wild type
312 protein (Figure 5C, right-hand panel), confirming that the efficiency with which SPP cleaves
313 the UL40 signal sequence does affect the resulting up-regulation of HLA-E.

314 *The mature UL40 and Rh67 proteins reduce surface expression of MHC class I proteins*
315 No function has been ascribed to the mature UL40 protein, and the lack of significant
316 sequence homology means that it is not clear if the mature Rh67 protein is its functional
317 homologue. However, in the course of these studies, we observed a slight but reproducible
318 reduction in HLA-E expression when the HLA-E*01:03 SCD was co-expressed with UL40 or
319 Rh67 with the HLA-E signal sequence (ESS-UL40 or ESS-Rh67, respectively – Figure 5A). A
320 more marked reduction in surface expression of HLA-E by ESS-UL40 and ESS-Rh67 was seen
321 when expression of the HLA-E*01:01 SCD was increased by co-expression of the VL9 peptide
322 in the cytosol (Figure 5B). Importantly, this effect was not limited to HLA-E. Staining with
323 the pan-MHC class I antibody W6/32 showed that expression of wild type UL40 or Rh67
324 (Figure 5C), or the signal sequence swapped versions (Figure S7A), resulted in a reduction in
325 the cell surface expression of MHC Ia that was comparable to that induced by the HCMV
326 US3 and US6 proteins (Figure S7B).

327 This suggests that the mature UL40 and Rh67 proteins represent previously undescribed
328 components of the HCMV and RhCMV immune evasion arsenal, although the effect on HLA-
329 E expression would appear to be at odds with the function of the UL40 and Rh67 signal
330 sequences. However, the increase in HLA-E expression mediated by the natural signal
331 sequences of UL40 and Rh67 is not dependent on TAP, in contrast to the expression of HLA-
332 E in the experiments shown in Figure 5A and B. Therefore, we investigated if the mature
333 UL40 and Rh67 proteins are also able to counteract the up-regulation of HLA-E expression
334 mediated by their signal sequence VL9 peptides.

335 Surface expression of the HLA-E*01:01 SCD in the presence of the UL40 signal sequence
336 (linked to HLa-A*02, USS-A2) was reduced by co-expression of either ESS-UL40 or ESS-Rh67
337 (Figure 6A), although HLA-E expression was still higher than when the HLA-E SCD was
338 expressed alone. This suggests that the level of expression of HLA-E in HCMV-infected cells
339 represents a balance between the competing processes mediated by the UL40 signal
340 sequence and the mature UL40 protein. In contrast, the up-regulation of HLA-E expression
341 caused by the Rh67 signal sequence (when RSS-A2 was co-expressed) was unaffected by
342 expression of the mature UL40 or Rh67 proteins (Figure 6B). At present, it is not clear what
343 underpins this difference in the sensitivity of the UL40 and Rh67 VL9 peptides to inhibition
344 by the mature UL40 and Rh67 proteins, and it may suggest that expression of Mamu-E on
345 RhCMV-infected cells may be higher than that of HLA-E on HCMV-infected cells. (2287
346 words)

347 DISCUSSION

348 We previously showed that Rh67 was likely to be the RhCMV functional homologue of the
349 HCMV UL40 protein (19), a finding strengthened by the recent demonstration that the up-

350 regulation of MHC-E expression mediated by Rh67 does not require TAP (20). We have now
351 extended our characterisation of Rh67 to show striking functional similarity with UL40,
352 despite the sequence of the two proteins showing little sequence homology. In particular,
353 the architecture of the signal sequences appears identical: both possess N-terminal
354 sequences that are essential for the up-regulation of HLA-E, and that confer TAP-
355 independence on the HLA-A*02 VL9 peptide when placed at the start of the HLA-A*02 signal
356 sequence. Moreover, despite the differences in the two signal sequences, processing of
357 both the UL40 and Rh67 VL9 peptides also depends on cleavage by the host protease SPP.
358 Most strikingly, however, our results also show that the mature UL40 and Rh67 proteins
359 contribute to the broader immune evasion orchestrated by CMV by decreasing the surface
360 expression of classical MHC class I proteins. At present, we have only studied the effects of
361 the mature Rh67 protein on the expression of human MHC class I proteins, and it will be
362 important to confirm that Rh67 also down-regulates Mamu-A and Mamu-B.

363 Unexpectedly, we also observed that the mature UL40 and Rh67 proteins can both reduce
364 the surface expression of HLA-E. For UL40, this effect on HLA-E appears to extends to the
365 increase in HLA-E expression mediated by its own signal sequence VL9 peptide, while the
366 Rh67 VL9 peptide is unaffected by the action of the mature Rh67 or UL40 proteins. The
367 UL40 signal sequence is also processed less efficiently by SPP than the Rh67 signal sequence,
368 and this may again reduce the magnitude of HLA-E up-regulation mediated by UL40. This
369 inefficient cleavage may simply be an indirect effect of the UL40 signal sequence being a
370 poorer signal sequence than that of Rh67 (at least as measured by the surface expression of
371 HLA-A*02 when its own signal sequence is replaced by those of the CMV proteins –Figure
372 S2), resulting in a reduced amount of UL40 signal sequence available for SPP to process. The
373 mutation that improves the UL40 signal sequence as a substrate for SPP (Figure 4) has little

374 effect on its functioning as a signal sequence (Figure S2), however, suggesting that the
375 reduced processing by SPP does result from the UL40 signal sequence being a poor SPP
376 substrate.

377 One explanation for these differences is that UL40 may have evolved to limit the expression
378 of HLA-E on the surface of infected cells. Perhaps related to this, the most potent peptide
379 ligand for NKG2/CD94 interaction (VMAPRTLFL, from the HLA-G signal sequence) is rarely
380 found in the UL40 signal sequence (33), in contrast to the other MHC Ia VL9 variants (such as
381 VMPARTLLL, VMAPRTLIL, VMAPRTLVL, and VMAPRTVLL). Limiting HLA-E expression may be
382 necessary to prevent recognition by NK cells expressing the activating NKG2C/CD94 ligand,
383 which has a lower affinity for HLA-E than the inhibitory NKG2A/CD94 complex (34). The fact
384 that Rh67 does not appear to limit its own VL9 suggests that Mamu-E expression by RhCMV-
385 infected cells may be higher than HLA-E expression by HCMV-infected cells, perhaps
386 reflecting differences in immune surveillance between human and macaque NK cells and the
387 expression levels required to trigger CD94/NKG2 signalling.

388 At present it is not clear why the mature UL40 and Rh67 proteins can counteract the VL9
389 peptide of UL40 but not that of Rh67, nor how both peptides reach the lumen of the ER in
390 the absence of TAP. The TAP-dependence of the MHC Ia VL9 peptides results from SPP
391 cleaving the signal sequence downstream of the VL9 peptide, releasing an extended version
392 of the peptide back into the cytosol (3). The simplest mechanism for the TAP-independence
393 of the UL40 and Rh67 VL9 peptides, therefore, would be for SPP to cleave upstream of VL9,
394 releasing the precursor directly into the lumen of the ER. Prediction of SPP cleavage sites is
395 complicated by the lack of a consensus sequence, although it is likely that helix-breaking or
396 polar residues following a stretch of alpha helix are critical (35). While the N-terminus of

397 the Rh67 is sufficiently hydrophobic to tolerate being embedded in the ER membrane, as
398 required for cleavage by SPP, that of UL40 is not (containing more polar residues). It was
399 previously proposed that the N-terminus of the UL40 signal sequence would initially project
400 into the cytosol, undergoing a topological change to allow cleavage by SPP (16). No direct
401 evidence in support of this model was presented, and it required the UL40 signal sequence
402 adopting a type I orientation, while SPP cleaves proteins in the type II orientation (36).

403 Sequences similar to that thought to be cleaved by SPP in the HLA-A*02 signal sequence
404 (LLSG, immediately following the VL9 peptide) lie both upstream (LLSV) and downstream
405 (LLCT) of the Rh67 VL9 peptide, although the first of these is present immediately after the
406 start codon and may not be suitably positioned for cleavage by SPP. For UL40, the only
407 match in the signal sequence (LLCM) lies downstream of the VL9 peptide. Therefore, for
408 both signal sequences, it seems likely that SPP cleaves downstream of the VL9 peptide,
409 which would release the VL9 precursors back into the cytosol. The first 5 amino acids of the
410 Rh67 signal sequence are identical to those of a known TAP-independent epitope
411 (MLLVSPLLL) which corresponds to the first 10 amino acids of the calreticulin signal
412 sequence (37). This signal sequence can be cleaved by SPP (31), which would release a
413 fragment containing this epitope into the cytosol. It has been proposed this peptide can
414 enter the ER in the absence of TAP by virtue of its extremely hydrophobic sequence (16).
415 The fragment generated by SPP cleaving after the Rh67 VL9 peptide would also be
416 extremely hydrophobic (perhaps MLLSVAMVMAPRTLLLVV) and may behave in the same
417 way as the calreticulin peptide. In contrast, the longer peptide that would be generated by
418 cleavage of the UL40 signal sequence after the VL9 peptide would be significantly less
419 hydrophobic, and presumably enters the lumen of the ER by a different mechanism,

420 explaining the sensitivity of the UL40 VL9 peptide to the inhibitory effect of the mature
421 UL40 and Rh67 proteins.

422 Having shown a requirement for SPP cleavage, we are also interested in determining what
423 other host factors are required to process the mature VL9 peptide from both UL40 and
424 Rh67. If precursors of these peptides do return to the cytosol, they may require processing
425 at the C-terminus by the proteasome, as has been observed for the VL9 peptides from
426 classical MHC class I alleles (5). Removal of the amino acids upstream of the Rh67 and UL40
427 VL9 peptides is also presumably required, which may suggest the involvement of the ER-
428 resident aminopeptidases ERAP1 or ERAP2. Formation of the MHC Ia VL9 peptide also
429 requires removal of two amino acids from the N-terminus of the signal sequence, but the
430 enzyme responsible has yet to be identified. Disruption of ERAAP (the homologue of ERAP1)
431 in mice has been shown to alter the peptide presented by Qa-1, the murine homologue of
432 HLA-E (38). Rather than presenting the Qdm epitope (AMAPRTLLL) from the murine class
433 leader sequence, ERAAP-deficient cells present a peptide derived from the Fam49b protein
434 (FYAEATPML). However, it is not clear if this results from the absence of ERAAP preventing
435 correct processing of the Qdm peptide, or the Fam49b peptide no longer being destroyed.
436 Two different ERAP1 inhibitors have been shown to have no effect on surface expression of
437 HLA-E (5), suggesting that ERAP1 may not be responsible for trimming the class I VL9
438 peptide in humans. A role for ERAP1 may be unlikely for the UL40 VL9 peptide, however, as
439 a microRNA encoded in the HCMV US4 gene has been shown to downregulate expression of
440 one of the two known isoforms of ERAP1 (39), although no equivalent microRNA has been
441 described for RhCMV.

442 The level of sequence conservation between the mature UL40 and Rh67 proteins is similar
443 to that of the other HCMV/RhCMV immunomodulatory proteins (US2/Rh182, US3/Rh184,
444 US6/Rh185, US8/Rh186, US10/Rh187, and US11.Rh189), which show 21–30% sequence
445 identity and 33–43% sequence similarity (17). Both proteins are known to localise in the ER
446 (20), but it is not clear if they are associated with the membrane or present in the lumen.
447 The TMHMM webserver (40) predicts a single transmembrane domain for Rh67, but not for
448 UL40, and homology scanning using InterProScan (41) fails to identify any known functional
449 domains. The structures of these proteins predicted by AlphaFold2 using the ColabFold
450 server (42) also differ (Figure S8). The 5 models generated for UL40 show little similarity to
451 each other when aligned, whereas those for Rh67 are much more homogeneous, differing
452 primarily in the orientation of the C-terminus, which is predicted to form an extended alpha
453 helix in all 5 models. Therefore, determining if these proteins are indeed functional
454 homologues will require identification of the step(s) of the antigen processing and
455 presentation pathway that they target.

456 Evasion of NK cells by the maintenance of MHC-E expression during CMV infection appears
457 to be common through-out primate evolution, and VL9 peptides are found in all but one of
458 the known non-human primate (NHP) CMV genomes (Table S8). As with UL40 and Rh67,
459 the viral proteins that contain the peptide show very little sequence homology (Figure S9).
460 In each case, the peptide is predicted to be present in a signal sequence, and its position
461 reflects that of Rh67, with only the HCMV UL40 protein having an extended N-terminus. It
462 seems likely that these NHP proteins will also mediate TAP-independent up-regulation of
463 MHC-E expression, and the associated mature proteins may also play a role in decreasing
464 surface expression of MHC class I proteins. The exception to this is the peptide present in
465 chimpanzee CMV (TMAPKTLLL), which differs from that present in all the other NHP CMVs at

466 two positions. The position 1 threonine is unlikely to be significant for binding by MHC-E
467 (Figure S10), although it could affect processing of the peptide from the signal sequence,
468 but the lysine at position 5 is striking. Interaction of HLA-E with CD94/NKG2A is abrogated
469 by mutating position 5 of the VL9 peptide from arginine to lysine (43). Chimpanzees are the
470 closest living relatives of humans, and their CD94 and NKG2A genes show >96% sequence
471 identity to their human counterparts (44), and all Patr-A alleles have the canonical VL9
472 peptide in their leader sequences. This suggests that monitoring of class I expression by NK
473 cells via MHC-E/CD94/NKG2A in chimpanzees will mirror that of humans. The chimpanzee
474 CMV UL40 protein is also the most closely related to HCMV UL40 of all the NHP orthologues,
475 but the sequence upstream of the TL9 peptide in chimpanzee CMV is not as long as that of
476 HCMV, and not as hydrophobic as that in the other NHP CMVs. Given the extensive in vitro
477 culturing of this stain prior to sequencing (45), it will be interesting to determine if this
478 sequence is truly representative of chimpanzee CMV UL40, or if it reflects differences in how
479 chimpanzee CMV evades detection by NK cells.

480 In conclusion, we show that despite the similarities (including the location of the critical
481 sequences and processing of their VL9 peptides being dependent on cleavage by SPP) in
482 how the signal sequences of UL40 and Rh67 up-regulate HLA-E expression, there are
483 important differences. Processing of the UL40 VL9 peptide is less efficient than that of
484 Rh67, and the resulting increase in HLA-E expression appears to be partially counteracted by
485 the ability of the mature UL40 protein to inhibit surface expression of MHC class I. The
486 mature Rh67 protein is also able to inhibit MHC class I expression, but does not affect the
487 up-regulation of HLA-E mediated by its own signal sequence VL9 peptide. Such differences
488 may have important implications for the ability of a HCMV-vectored vaccine to induce

489 protective HLA-E-restricted responses (analogous to the protective Mamu-E responses
490 elicited by the RvCMV 68.1 SIV vaccine) and warrants further investigation.

491 ACKNOWLEDGEMENTS

492 We thank Andrea Margi (NDMRB, University of Oxford) for the kind gift of the HCV Core-E1
493 expression plasmid. Andrew Worth of the Jenner Institute for cell sorting, and Tim Rostron
494 and John Frankland of the Weatherall Institute of Molecular Medicine Sequencing Facility
495 for plasmid sequencing. We also thank Geraldine Gillespie and past and present members
496 of the Gillespie, Borrow and McMichael laboratories for helpful discussions while this work
497 was being undertaken. The work was supported by funding from the Bill and Melinda Gates
498 Foundation (OPP1133649) and the Chinese Academy of Medical Sciences (CAMS) Innovation
499 Fund for Medical Sciences (CIFMS 2018-I2M-2-002). P.B. and A.J.M. are Jenner Institute
500 Investigators.

501 S.B. conceived and performed the experiments and analysed the data. P.B. and A.J.M.
502 supervised the work and analysed data. S.B. wrote the manuscript, which was reviewed,
503 edited and approved by all authors. The authors declare that they have no conflicts of
504 interest with respect to the publication of this manuscript.

505 REFERENCES

506 1. Braud V, Jones EY, McMichael A. 1997. The human major histocompatibility complex
507 class Ib molecule HLA-E binds signal sequence-derived peptides with primary anchor
508 residues at positions 2 and 9. Eur J Immunol 27:1164-9.

509 2. Lee N, Goodlett DR, Ishitani A, Marquardt H, Geraghty DE. 1998. HLA-E surface
510 expression depends on binding of TAP-dependent peptides derived from certain HLA
511 class I signal sequences. J Immunol 160:4951-60.

512 3. Lemberg MK, Bland FA, Weihofen A, Braud VM, Martoglio B. 2001. Intramembrane
513 proteolysis of signal peptides: an essential step in the generation of HLA-E epitopes. J
514 Immunol 167:6441-6.

515 4. Weihofen A, Lemberg MK, Ploegh HL, Bogyo M, Martoglio B. 2000. Release of signal
516 peptide fragments into the cytosol requires cleavage in the transmembrane region by a
517 protease activity that is specifically blocked by a novel cysteine protease inhibitor. J Biol
518 Chem 275:30951-6.

519 5. Bland FA, Lemberg MK, McMichael AJ, Martoglio B, Braud VM. 2003. Requirement of
520 the proteasome for the trimming of signal peptide-derived epitopes presented by the
521 nonclassical major histocompatibility complex class I molecule HLA-E. J Biol Chem
522 278:33747-52.

523 6. Braud VM, Allan DS, Wilson D, McMichael AJ. 1998. TAP- and tapasin-dependent HLA-E
524 surface expression correlates with the binding of an MHC class I leader peptide. Curr
525 Biol 8:1-10.

526 7. Braud VM, Allan DS, O'Callaghan CA, Soderstrom K, D'Andrea A, Ogg GS, Lazetic S,
527 Young NT, Bell JI, Phillips JH, Lanier LL, McMichael AJ. 1998. HLA-E binds to natural killer
528 cell receptors CD94/NKG2A, B and C. Nature 391:795-9.

529 8. Lee N, Llano M, Carretero M, Ishitani A, Navarro F, Lopez-Botet M, Geraghty DE. 1998.
530 HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. Proc Natl
531 Acad Sci U S A 95:5199-204.

532 9. Borrego F, Alonso MC, Galiani MD, Carracedo J, Ramirez R, Ostos B, Pena J, Solana R.
533 1999. NK phenotypic markers and IL2 response in NK cells from elderly people. Exp
534 Gerontol 34:253-65.

535 10. Furman MH, Dey N, Tortorella D, Ploegh HL. 2002. The human cytomegalovirus US10
536 gene product delays trafficking of major histocompatibility complex class I molecules. J
537 Virol 76:11753-6.

538 11. Wiertz EJ, Jones TR, Sun L, Bogyo M, Geuze HJ, Ploegh HL. 1996. The human
539 cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the
540 endoplasmic reticulum to the cytosol. Cell 84:769-79.

541 12. Wiertz EJ, Tortorella D, Bogyo M, Yu J, Mothes W, Jones TR, Rapoport TA, Ploegh HL.
542 1996. Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum
543 to the proteasome for destruction. Nature 384:432-8.

544 13. Ahn K, Gruhler A, Galocha B, Jones TR, Wiertz EJ, Ploegh HL, Peterson PA, Yang Y, Fruh
545 K. 1997. The ER-luminal domain of the HCMV glycoprotein US6 inhibits peptide
546 translocation by TAP. Immunity 6:613-21.

547 14. Park B, Kim Y, Shin J, Lee S, Cho K, Fruh K, Lee S, Ahn K. 2004. Human cytomegalovirus
548 inhibits tapasin-dependent peptide loading and optimization of the MHC class I peptide
549 cargo for immune evasion. Immunity 20:71-85.

550 15. Tomasec P, Braud VM, Rickards C, Powell MB, McSharry BP, Gadola S, Cerundolo V,
551 Borysiewicz LK, McMichael AJ, Wilkinson GW. 2000. Surface expression of HLA-E, an
552 inhibitor of natural killer cells, enhanced by human cytomegalovirus gpUL40. Science
553 287:1031.

554 16. Prod'homme V, Tomasec P, Cunningham C, Lemberg MK, Stanton RJ, McSharry BP,
555 Wang EC, Cuff S, Martoglio B, Davison AJ, Braud VM, Wilkinson GW. 2012. Human
556 cytomegalovirus UL40 signal peptide regulates cell surface expression of the NK cell
557 ligands HLA-E and gpUL18. J Immunol 188:2794-804.

558 17. Pande NT, Powers C, Ahn K, Fruh K. 2005. Rhesus cytomegalovirus contains functional
559 homologues of US2, US3, US6, and US11. J Virol 79:5786-98.

560 18. Richards R, Scholz I, Powers C, Skach WR, Fruh K. 2011. The cytoplasmic domain of
561 rhesus cytomegalovirus Rh178 interrupts translation of major histocompatibility class I
562 leader peptide-containing proteins prior to translocation. J Virol 85:8766-76.

563 19. Hansen SG, Wu HL, Burwitz BJ, Hughes CM, Hammond KB, Ventura AB, Reed JS, Gilbride
564 RM, Ainslie E, Morrow DW, Ford JC, Selseth AN, Pathak R, Malouli D, Legasse AW,
565 Axthelm MK, Nelson JA, Gillespie GM, Walters LC, Brackenridge S, Sharpe HR, Lopez CA,
566 Fruh K, Korber BT, McMichael AJ, Gnanakaran S, Sacha JB, Picker LJ. 2016. Broadly
567 targeted CD8(+) T cell responses restricted by major histocompatibility complex E.
568 Science 351:714-20.

569 20. Verweij MC, Hansen SG, Iyer R, John N, Malouli D, Morrow D, Scholz I, Womack J,
570 Abdulhaqq S, Gilbride RM, Hughes CM, Ventura AB, Ford JC, Selseth AN, Oswald K,
571 Shoemaker R, Berkemeier B, Bosche WJ, Hull M, Shao J, Sacha JB, Axthelm MK, Edlefsen
572 PT, Lifson JD, Picker LJ, Fruh K. 2021. Modulation of MHC-E transport by viral decoy
573 ligands is required for RhCMV/SIV vaccine efficacy. Science 372.

574 21. Hansen SG, Ford JC, Lewis MS, Ventura AB, Hughes CM, Coyne-Johnson L, Whizin N,
575 Oswald K, Shoemaker R, Swanson T, Legasse AW, Chiuchiolo MJ, Parks CL, Axthelm MK,
576 Nelson JA, Jarvis MA, Piatak M, Jr., Lifson JD, Picker LJ. 2011. Profound early control of
577 highly pathogenic SIV by an effector memory T-cell vaccine. Nature 473:523-7.

578 22. Hansen SG, Piatak M, Jr., Ventura AB, Hughes CM, Gilbride RM, Ford JC, Oswald K,
579 Shoemaker R, Li Y, Lewis MS, Gilliam AN, Xu G, Whizin N, Burwitz BJ, Planer SL, Turner
580 JM, Legasse AW, Axthelm MK, Nelson JA, Fruh K, Sacha JB, Estes JD, Keele BF, Edlefsen
581 PT, Lifson JD, Picker LJ. 2013. Immune clearance of highly pathogenic SIV infection.
582 Nature 502:100-4.

583 23. Malouli D, Hansen SG, Hancock MH, Hughes CM, Ford JC, Gilbride RM, Ventura AB,
584 Morrow D, Randall KT, Taher H, Uebelhoer LS, McArdle MR, Papen CR, Espinosa
585 Trethewy R, Oswald K, Shoemaker R, Berkemeier B, Bosche WJ, Hull M, Greene JM,
586 Axthelm MK, Shao J, Edlefsen PT, Grey F, Nelson JA, Lifson JD, Streblow D, Sacha JB,
587 Fruh K, Picker LJ. 2021. Cytomegaloviral determinants of CD8(+) T cell programming and
588 RhCMV/SIV vaccine efficacy. Sci Immunol 6.

589 24. Hansen SG, Hancock MH, Malouli D, Marshall EE, Hughes CM, Randall KT, Morrow D,
590 Ford JC, Gilbride RM, Selseth AN, Trethewy RE, Bishop LM, Oswald K, Shoemaker R,
591 Berkemeier B, Bosche WJ, Hull M, Silipino L, Nekorchuk M, Busman-Sahay K, Estes JD,
592 Axthelm MK, Smedley J, Shao D, Edlefsen PT, Lifson JD, Fruh K, Nelson JA, Picker LJ.
593 2022. Myeloid cell tropism enables MHC-E-restricted CD8(+) T cell priming and vaccine
594 efficacy by the RhCMV/SIV vaccine. Sci Immunol 7:eabn9301.

595 25. Higuchi R, Krummel B, Saiki RK. 1988. A general method of in vitro preparation and
596 specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic
597 Acids Res 16:7351-67.

598 26. Wang T, Wei JJ, Sabatini DM, Lander ES. 2014. Genetic screens in human cells using the
599 CRISPR-Cas9 system. Science 343:80-4.

600 27. Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, Thakore PI,
601 Glass KA, Ousterout DG, Leong KW, Guilak F, Crawford GE, Reddy TE, Gersbach CA.
602 2013. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat
603 Methods 10:973-6.

604 28. Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K. 2013. Enhanced
605 efficiency of human pluripotent stem cell genome editing through replacing TALENs
606 with CRISPRs. Cell Stem Cell 12:393-4.

607 29. Brinkman EK, Chen T, Amendola M, van Steensel B. 2014. Easy quantitative assessment
608 of genome editing by sequence trace decomposition. Nucleic Acids Res 42:e168.

609 30. Strong RK, Holmes MA, Li P, Braun L, Lee N, Geraghty DE. 2003. HLA-E allelic variants.
610 Correlating differential expression, peptide affinities, crystal structures, and thermal
611 stabilities. J Biol Chem 278:5082-90.

612 31. Lemberg MK, Martoglio B. 2002. Requirements for signal peptide peptidase-catalyzed
613 intramembrane proteolysis. Mol Cell 10:735-44.

614 32. McLauchlan J, Lemberg MK, Hope G, Martoglio B. 2002. Intramembrane proteolysis
615 promotes trafficking of hepatitis C virus core protein to lipid droplets. EMBO J 21:3980-
616 8.

617 33. Hammer Q, Ruckert T, Borst EM, Dunst J, Haubner A, Durek P, Heinrich F, Gasparoni G,
618 Babic M, Tomic A, Pietra G, Nienen M, Blau IW, Hofmann J, Na IK, Prinz I, Koenecke C,
619 Hemmati P, Babel N, Arnold R, Walter J, Thurley K, Mashreghi MF, Messerle M,
620 Romagnani C. 2018. Peptide-specific recognition of human cytomegalovirus strains
621 controls adaptive natural killer cells. Nat Immunol 19:453-463.

622 34. Kaiser BK, Barahmand-Pour F, Paulsene W, Medley S, Geraghty DE, Strong RK. 2005.
623 Interactions between NKG2x immunoreceptors and HLA-E ligands display overlapping
624 affinities and thermodynamics. J Immunol 174:2878-84.

625 35. Lemberg MK, Martoglio B. 2004. On the mechanism of SPP-catalysed intramembrane
626 proteolysis; conformational control of peptide bond hydrolysis in the plane of the
627 membrane. FEBS Lett 564:213-8.

628 36. Mentrup T, Fluhrer R, Schroder B. 2017. Latest emerging functions of SPP/SPPL
629 intramembrane proteases. Eur J Cell Biol 96:372-382.

630 37. Henderson RA, Michel H, Sakaguchi K, Shabanowitz J, Appella E, Hunt DF, Engelhard VH.
631 1992. HLA-A2.1-associated peptides from a mutant cell line: a second pathway of
632 antigen presentation. Science 255:1264-6.

633 38. Nagarajan NA, Gonzalez F, Shastri N. 2012. Nonclassical MHC class Ib-restricted
634 cytotoxic T cells monitor antigen processing in the endoplasmic reticulum. Nat Immunol
635 13:579-86.

636 39. Kim S, Lee S, Shin J, Kim Y, Evnouchidou I, Kim D, Kim YK, Kim YE, Ahn JH, Riddell SR,
637 Stratikos E, Kim VN, Ahn K. 2011. Human cytomegalovirus microRNA miR-US4-1 inhibits
638 CD8(+) T cell responses by targeting the aminopeptidase ERAP1. Nat Immunol 12:984-
639 91.

640 40. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane
641 protein topology with a hidden Markov model: application to complete genomes. J Mol
642 Biol 305:567-80.

643 41. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell
644 A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R,
645 Hunter S. 2014. InterProScan 5: genome-scale protein function classification.
646 Bioinformatics 30:1236-40.

647 42. Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold:
648 making protein folding accessible to all. Nat Methods doi:10.1038/s41592-022-01488-1.

649 43. Miller JD, Weber DA, Ibegbu C, Pohl J, Altman JD, Jensen PE. 2003. Analysis of HLA-E
650 peptide-binding specificity and contact residues in bound peptide required for
651 recognition by CD94/NKG2. J Immunol 171:1369-75.

652 44. Shum BP, Flodin LR, Muir DG, Rajalingam R, Khakoo SI, Cleland S, Guethlein LA, Uhrberg
653 M, Parham P. 2002. Conservation and variation in human and common chimpanzee
654 CD94 and NKG2 genes. J Immunol 168:240-52.

655 45. Davison AJ, Dolan A, Akter P, Addison C, Dargan DJ, Alcendor DJ, McGeoch DJ, Hayward
656 GS. 2003. The human cytomegalovirus genome revisited: comparison with the
657 chimpanzee cytomegalovirus genome. J Gen Virol 84:17-28.

658 46. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN,
659 Potter SC, Finn RD, Lopez R. 2019. The EMBL-EBI search and sequence analysis tools
660 APIs in 2019. Nucleic Acids Res 47:W636-W641.

661

662 # FIGURE LEGENDS

663 *FIGURE 1: UL40 and Rh67 increase HLA-E expression by a TAP-independent pathway in*
664 *293T cells*

665 **(A)** Alignment of the sequences of the HCMV UL40 gene (strain Toledo) and the Rhesus CMV
666 Rh67 gene (strain 68-1). The location of the VL9 peptides, the only region of significant
667 sequence similarity shared by the two proteins, is indicated. The sequences were aligned
668 using EMBOSS Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) (46), and amino
669 acids are coloured according to the standard RasMol scheme. **(B, C)** Representative HLA-E
670 expression in 293T cells (panel B), or the TAP-deficient 293T clones TAP1-1 or TAP 2-4 (panel
671 C), co-transfected with the HLA-E*01:01 SCD expression plasmid and plasmids expressing an
672 irrelevant protein (filled light grey histogram), HLA-A*02 (orange line), UL40 (purple line), or
673 Rh67 (pink line). Mock transfected cells are shown as the filled dark grey histogram in panel
674 B. The sequence alignment in panel B shows a comparison of the signal sequences of the
675 transfected UL40, Rh67 and HLA-A*02 proteins.

676

677 *FIGURE 2: The N-terminus of the Rh67 signal sequence is critical for TAP-independent*
678 *up-regulation of HLA-E*

679 Representative HLA-E expression in 293T cells (panels **A** and **B**), or the TAP1-1 or TAP2-4
680 clones (panels **C** and **D**) co-transfected with the HLA-E*01:03 SCD expression plasmid and
681 plasmids expressing an irrelevant protein (light grey filled histogram), Rh67 with the HLA-E
682 signal sequences (ESS-Rh67; dashed pink line in panel A), HLA-A*02 with the Rh67 signal
683 sequence (RSS-A2; dashed orange line in panel B), Rh67 (pink line in panel C), Rh67 with the
684 N-terminal truncation shown in the sequence alignment (Rh67 ΔN; dotted pink line in panel
685 C), HLA-A*02 (orange line in panel D), or HLA-A*02 with the first two amino acids of the

686 signal sequence replaced with the first 7 amino acids of the Rh67 signal sequence (RVA-A2;
687 dotted orange line in panel D).

688

689 *FIGURE 3: Cleavage by SPP is essential for the up-regulation of HLA-E mediated by*
690 *UL40 and Rh67*

691 **(A)** Representative HLA-E expression in the four HM13 knockout CRISPR/Cas9 single cell
692 clones transfected with the HLA-E*01:03 SCD expression plasmid and plasmids expressing
693 an irrelevant protein (light grey filled histogram), UL40 (purple line), or Rh67 (pink line). **(B)**
694 Processing of the HCV Core protein, assessed by western blotting using an antibody specific
695 for the HCV Core protein, in the parental 293T cells and all four HM13 CRISPR/Cas9 clones
696 transfected with a plasmid expressing the Core-E1 precursor. **(C)** Representative HLA-E
697 expression in clone HM13-3 transfected with the HLA-E*01:03 SCD expression plasmid,
698 plasmids expressing UL40 (left-hand column, purple) or Rh67 (right-hand column, pink), as
699 well as plasmids expressing either an irrelevant protein (filled purple or pink histograms
700 histogram), or wild type SPP protein (purple or pink lines). Three different quantities of the
701 wild type SPP expression plasmid were used in panel E (375ng for the top row, 187ng for the
702 middle row, and 93ng for the bottom row, with the total amount of transfected DNA being
703 maintained by the addition of the required quantity of an irrelevant control plasmid). **(D)**
704 Processing of the HCV Core protein in clone HM13-3 transfected with a plasmid expressing
705 the Core-E1 precursor, as well as plasmids expressing an irrelevant protein (denoted by "-"),
706 wild type SPP, or a catalytically-inactive D219A SPP mutant. **(E)** Representative HLA-E
707 expression in clone HM13-3 transfected with the HLA-E*01:03 SCD expression plasmid,
708 plasmids expressing UL40 (left-hand column, purple) or Rh67 (right-hand column, pink), as

709 well as plasmids expressing either an irrelevant protein (filled purple or pink histograms

710 histogram), or the catalytically-inactive D219A SPP mutant (purple or pink lines).

711

712 *FIGURE 4: Inefficient processing of the UL40 signal sequence by SPP affects up-*
713 *regulation of HLA-E*

714 **(A)** Representative HLA-E expression in 293T cells transfected with the HLA-E*01:03 SCD

715 expression plasmid and plasmids expressing UL40 (left-hand graph, purple) or Rh67 (right-

716 hand graph, pink), as well as plasmids expressing an irrelevant protein (filled histogram) or

717 wild type SPP protein (coloured line). Mock transfected cells are shown as the filled dark

718 grey histograms. **(B)** Location of the double mutation (TT) that was previously suggested to

719 improve cleavage of the UL40 signal sequence by SPP. **(C)** Representative HLA-E expression

720 in clone HM13-1 (left-hand graph) or 293T cells (right-hand graph) transfected with the

721 ESCD expression plasmid and plasmids expressing an irrelevant protein (light grey filled

722 histograms), wild type UL40 (filled purple histogram), or the UL40 signal sequence mutant

723 (TT, red line).

724

725 *FIGURE 5: The mature UL40 and Rh67 proteins decrease surface expression of HLA-E*
726 *and classical MHC class I proteins*

727 **(A)** Representative HLA-E expression in 293T cells co-transfected with the HLA-E*01:03 SCD

728 expression plasmid and plasmids expressing an irrelevant protein (filled light grey

729 histogram), UL40 with the HLA-E signal sequence (ESS-UL40, purple line), or Rh67 with the

730 HLA-E signal sequence (ESS-Rh67, pink line). **(B)** Representative HLA-E expression in 293T

731 cells co-transfected with the HLA-E*01:01 SCD expression plasmid and a plasmid expressing

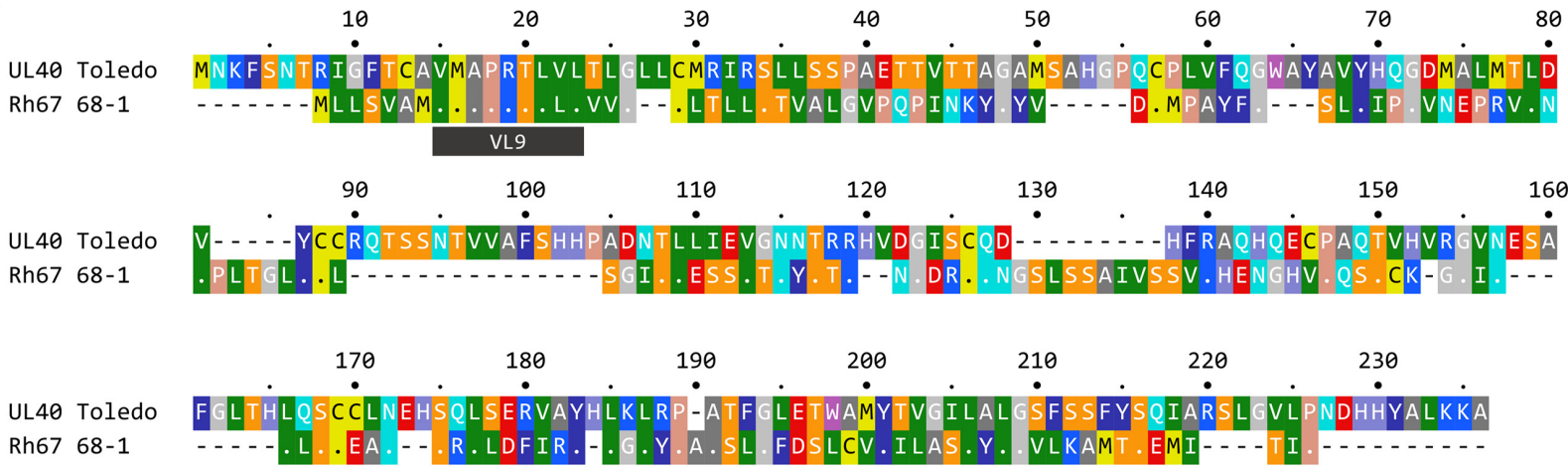
732 VL9 peptide in the cytosol, and plasmids expressing an irrelevant protein (blue filled

733 histogram), UL40 with the HLA-E signal sequence (ESS-UL40, purple line), or Rh67 with the
734 HLA-E signal sequence (ESS-Rh67, pink line). **(C)** MHC class I protein expression (as assessed
735 by W6/32 staining) of 293T cells transfected with plasmids expressing an irrelevant protein
736 (dark grey histogram), UL40 (purple line), or Rh67 (pink line).

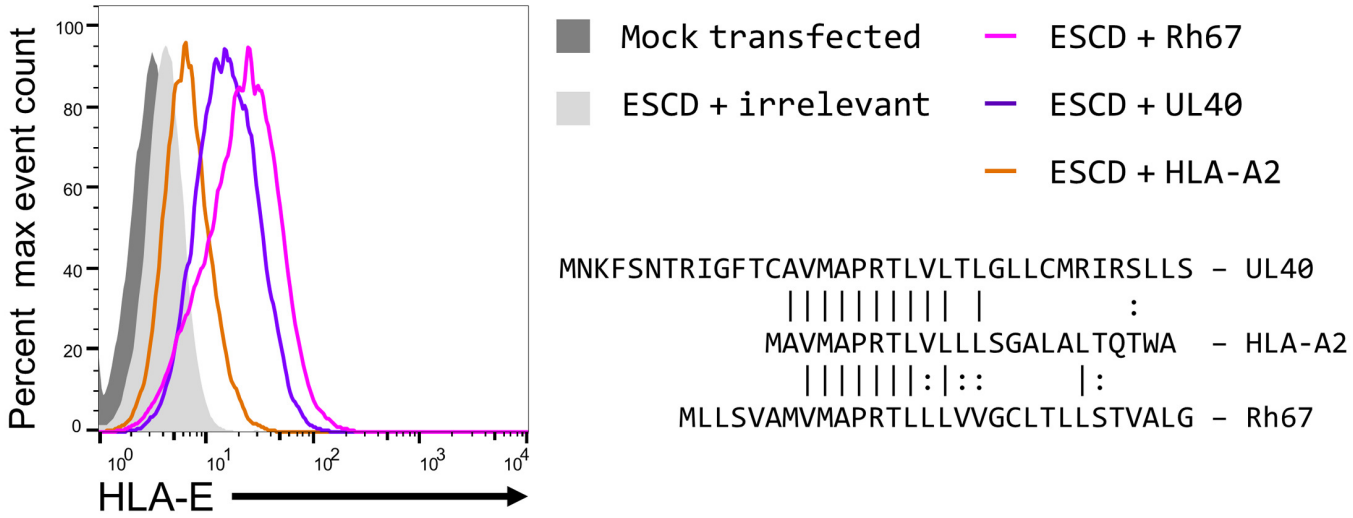
737

738 *FIGURE 6: The mature UL40 and Rh67 proteins can counteract the up-regulation of*
739 *HLA-E induced by the VL9 peptide of UL40 but not that of Rh67*
740 **(A)** Representative HLA-E expression in 293T cells co-transfected with the HLA-E*01:01 SCD
741 expression plasmid and plasmids expressing: i) only an irrelevant protein (light blue filled
742 histogram), ii) only HLA-A*02 with the UL40 signal sequence (USS-A2, filled purple
743 histogram), iii) HLA-A*02 with the UL40 signal sequence and UL40 with the HLA-E signal
744 sequence (USS-A2+ESS-UL40, purple line), or iv) HLA-A*02 with the UL40 signal sequence
745 and Rh67 with the HLA-E signal sequence (USS-A2+ESS-Rh67, pink line) **(B)** Representative
746 HLA-E expression in 293T cells co-transfected with the HLA-E*01:01 SCD expression plasmid
747 and plasmids expressing i) only an irrelevant protein (light blue filled histogram), ii) only
748 HLA-A*02 with the Rh67 signal sequence (RSS-A2, filled purple histogram), iii) HLA-A*02
749 with the Rh67 signal sequence and UL40 with the HLA-E signal sequence (RSS-A2+ESS-UL40,
750 purple line), or iv) HLA-A*02 with the Rh67 signal sequence and Rh67 with the HLA-E signal
751 sequence (RSS-A2+ESS-Rh67, pink line). In both panels, mock transfected cells are shown as
752 the filled dark grey histograms.

A


10 20 30 40 50 60 70 80
UL40 Toledo MNKFSNTRIGFTCAVMAPRTLVLTLGLLCMRIRSLLSSPAETTVTTAGAMSAHGPQCPLVFQGWAYAVYHQGDMALMTLD
Rh67 68-1 -------MLLSVAM.....L.VV.--.LTLL.TVALGVPQPINKY.YV-----D.MPAYF.---SL.IP.VNEPRV.N
VL9
90 100 110 120 130 140 150 160
UL40 Toledo V-----YCCRQTSSNTVVAFSHHPADNTLLIEVGNNTRRHVDGISCQD---------HFRAQHQECPAQTVHVRGVNESA
Rh67 68-1 .PLTGL..L----------SGI..ESS.T.Y.T.--N.DR..NGSLSSAIVSSV.HENGHV.QS.CK-G.I.---
170 180 190 200 210 220 230
UL40 Toledo FGLTHLQSCCLNEHSQLSERVAYHLKLRP-ATFGLETWAMYTVGILALGSFSSFYSQIARSLGVLPNDHHYALKKA
Rh67 68-1 -----.L..EA.--.R.LDFIR.-.G.Y.A.SL.FDSLCV.ILAS.Y..VLKAMT.EMI----TI.----------

B


Percent max event count
HLA-E
Mock transfected
ESCD + irrelevant
ESCD + Rh67
ESCD + UL40
ESCD + HLA-A2
MNKFSNTRIGFTCAVMAPRTLVLTLGLLCMRIRSLLS – UL40
MAVMAPRTLVLLLSGALALTQTWA – HLA-A2
MLLSVAMVMAPRTLLLVVGCLTLLSTVALG – Rh67

C

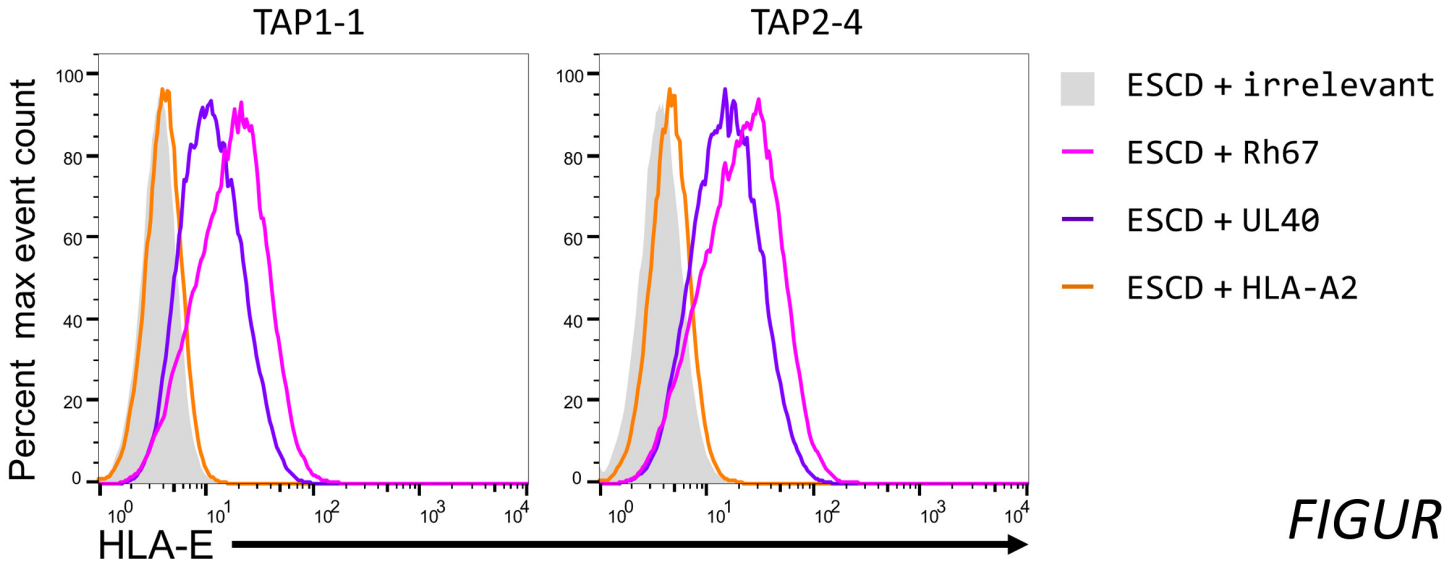

TAP1-1
TAP2-4
Percent max event count
HLA-E
ESCD + irrelevant
ESCD + Rh67
ESCD + UL40
ESCD + HLA-A2

FIGURE 1

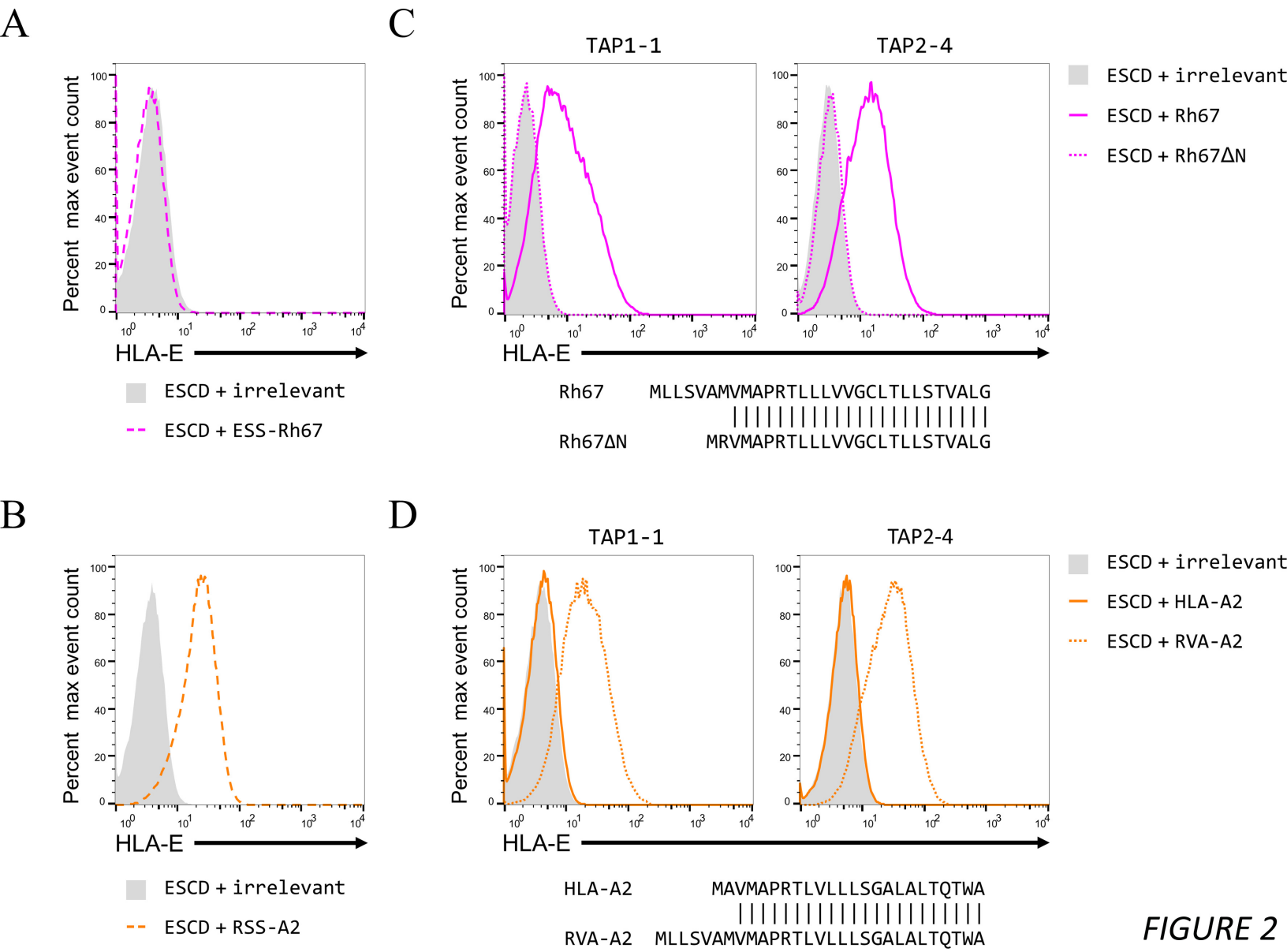
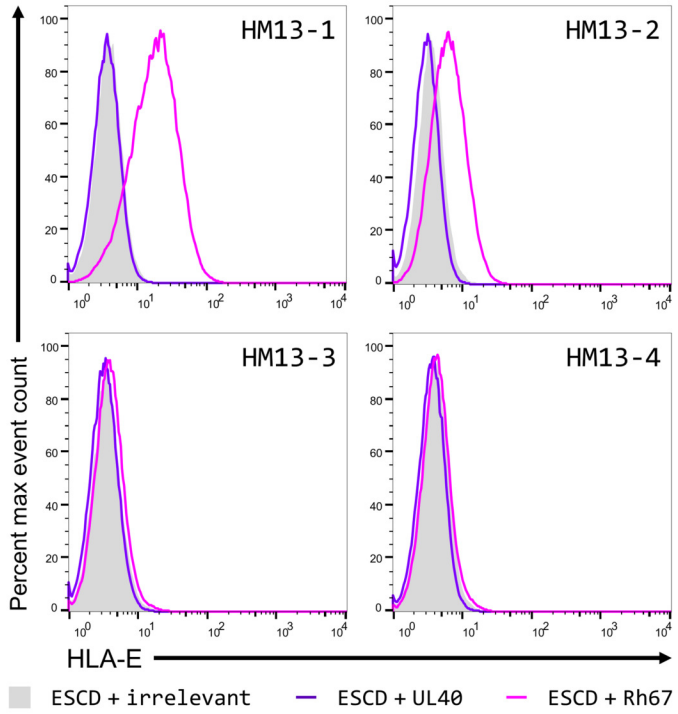
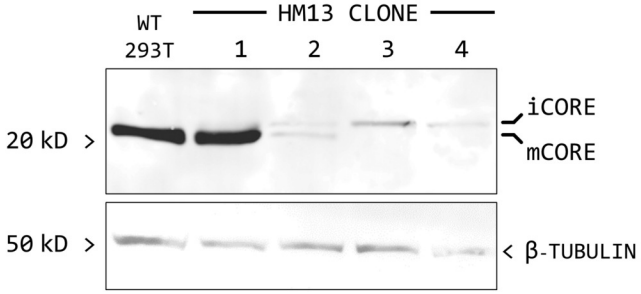

A
Percent max event count
HLA-E
ESCD + irrelevant
ESCD + ESS-Rh67
B
Percent max event count
HLA-E
ESCD + irrelevant
ESCD + RSS-A2
C
TAP1-1
TAP2-4
Percent max event count
HLA-E
ESCD + irrelevant
ESCD + Rh67
ESCD + Rh67ΔN
Rh67 MLLSVAMVMAPRTLLLVVGCLTLLSTVALG
Rh67ΔN MRVMAPRTLLLVVGCLTLLSTVALG
D
TAP1-1
TAP2-4
Percent max event count
HLA-E
ESCD + irrelevant
ESCD + HLA-A2
ESCD + RVA-A2
HLA-A2 MAVMAPRTLVLLLSGALALTQTWA
RVA-A2 MLLSVAMVMAPRTLVLLLSGALALTQTWA

FIGURE 2

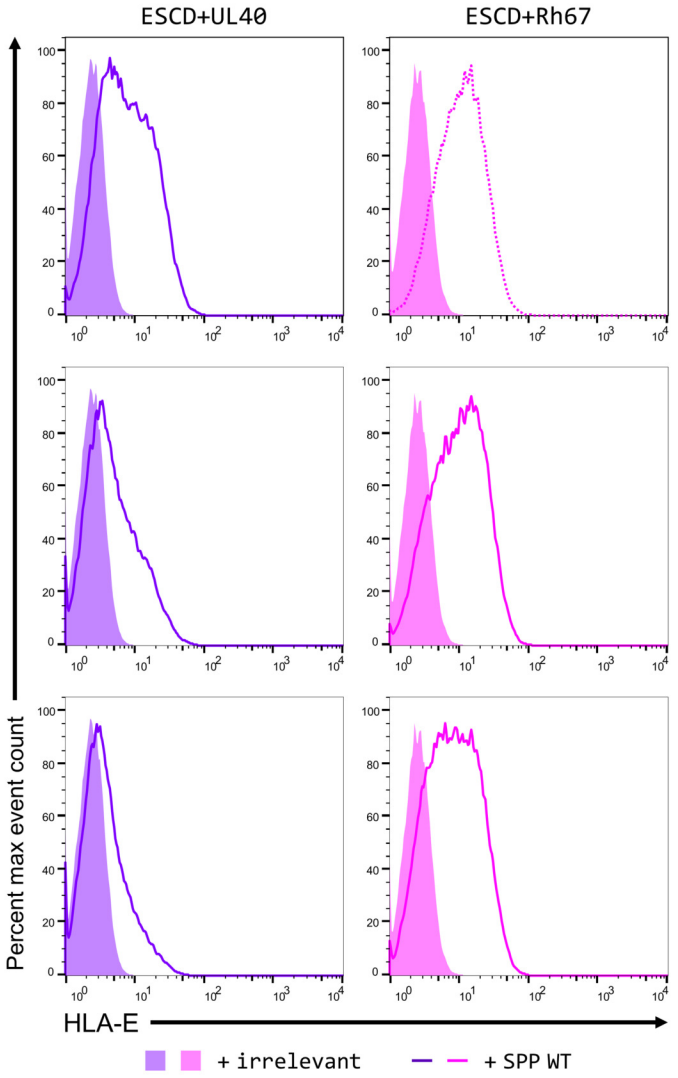
A



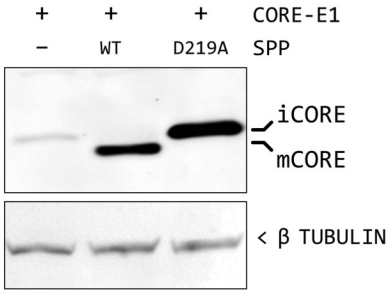
B



C



D



E

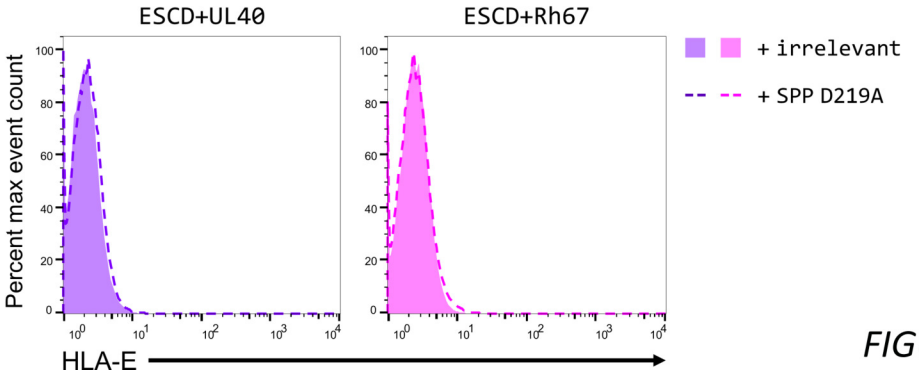
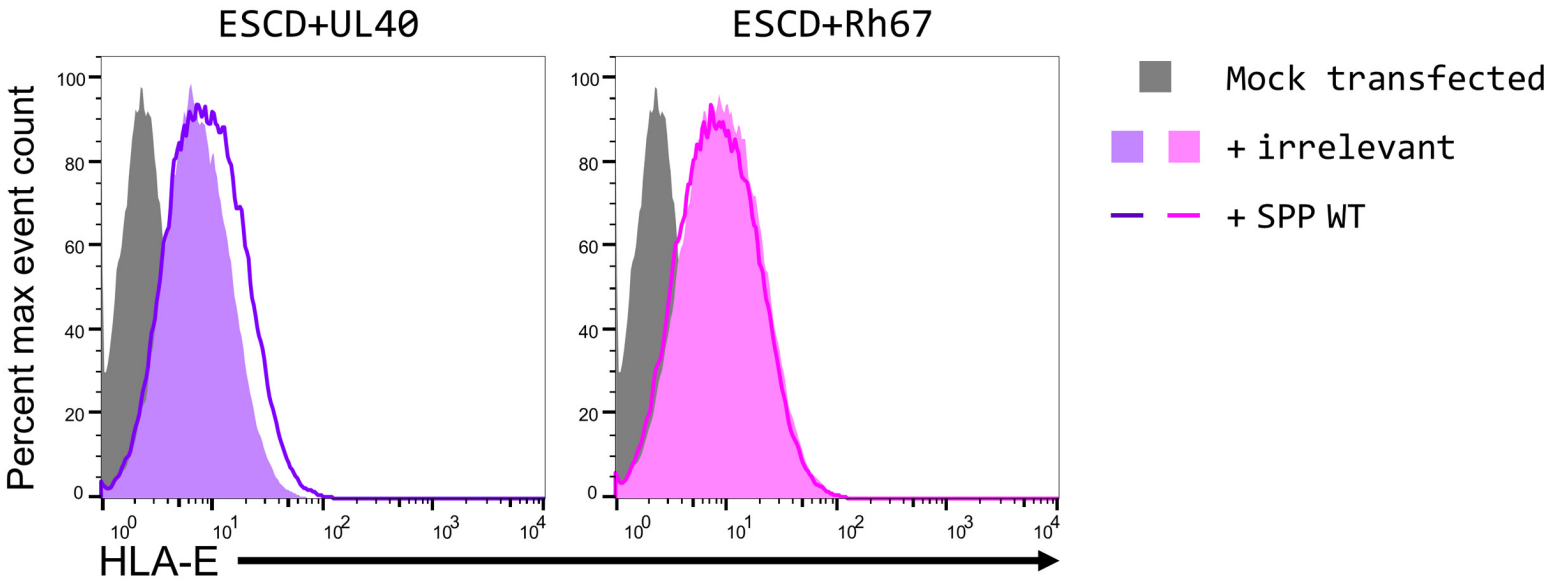


FIGURE 3

A



B

```
UL40 WT  MNKFSNTRIGFTCAVMAPRTLVLTLGLLCMRIRSLLS
UL40 TT  ..............................T.T....
```

C

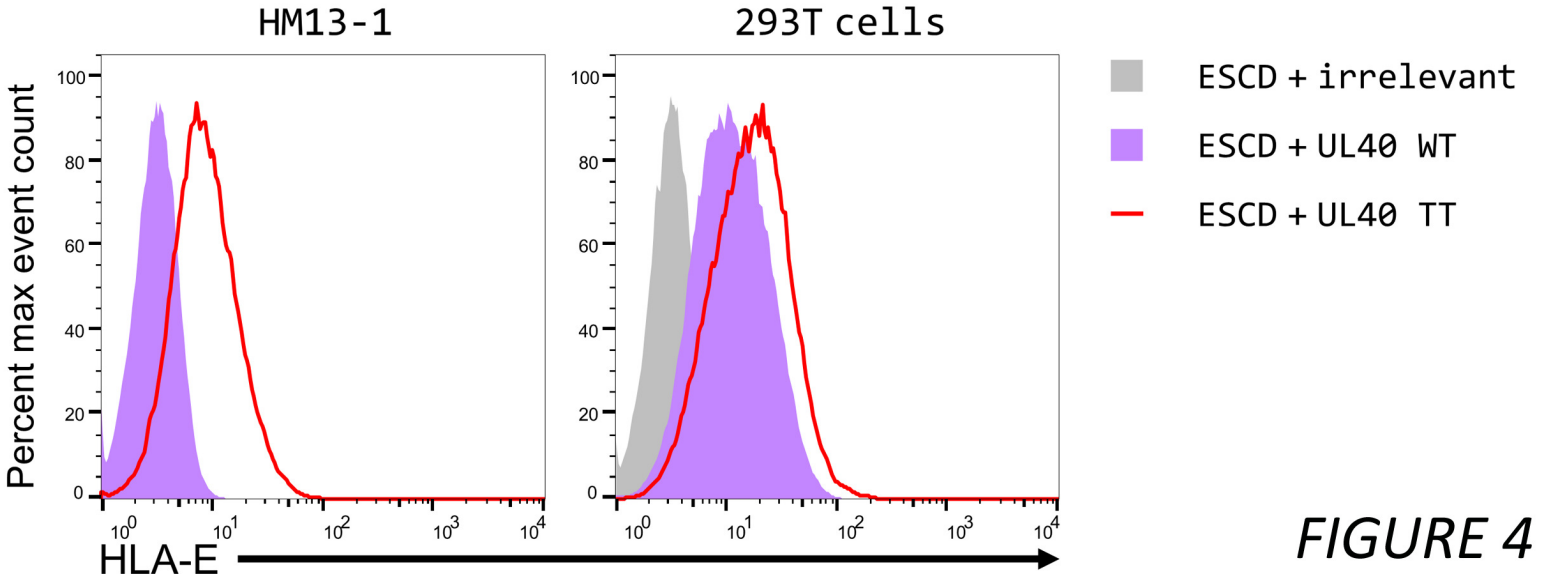
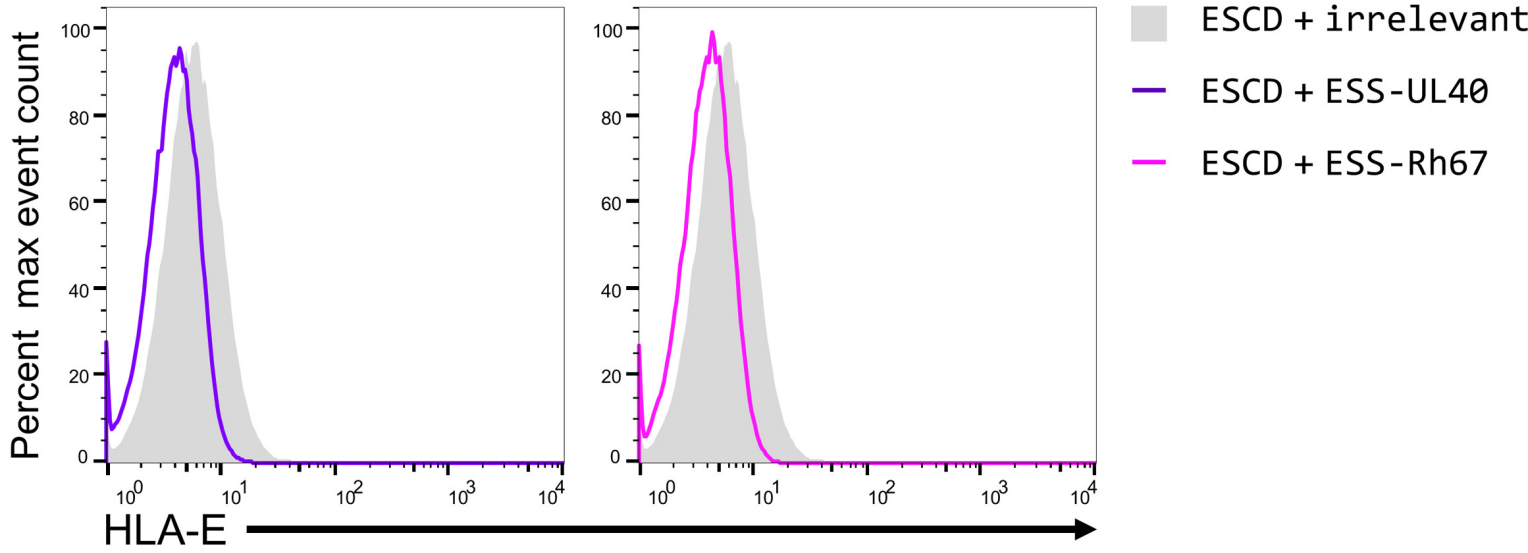
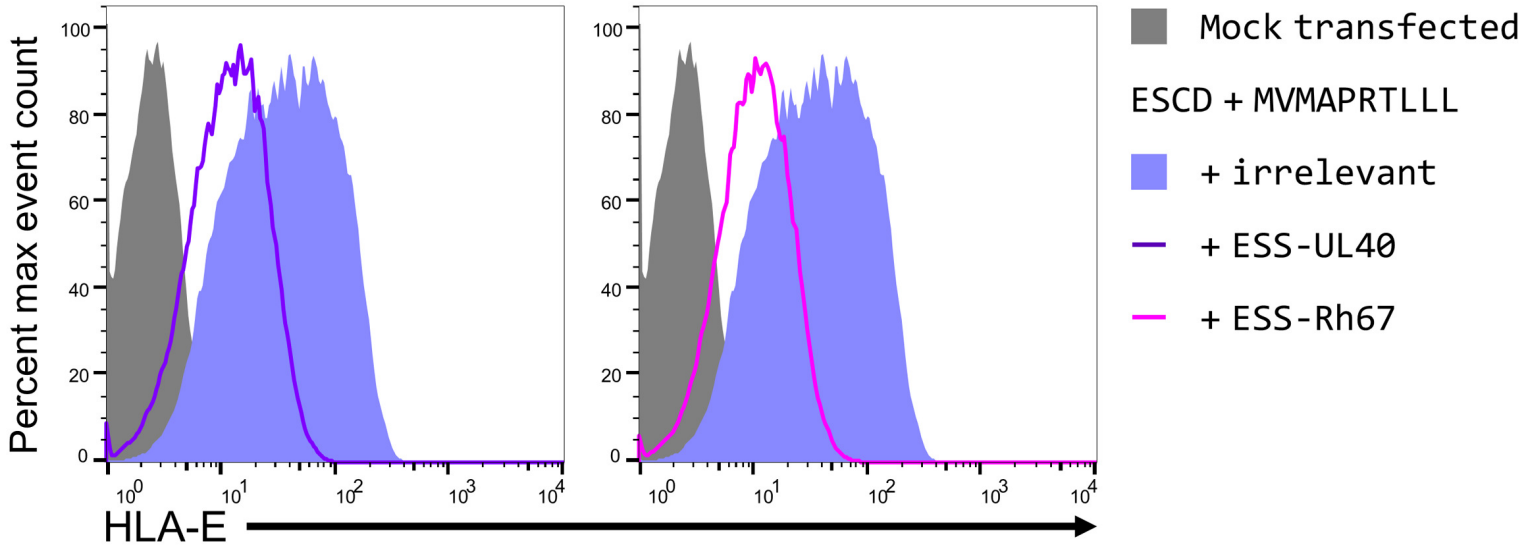


FIGURE 4

A


Percent max event count
HLA-E
ESCD + irrelevant
ESCD + ESS-UL40
ESCD + ESS-Rh67

B


Percent max event count
HLA-E
Mock transfected
ESCD + MVMAPRTLLL
+ irrelevant
+ ESS-UL40
+ ESS-Rh67

C

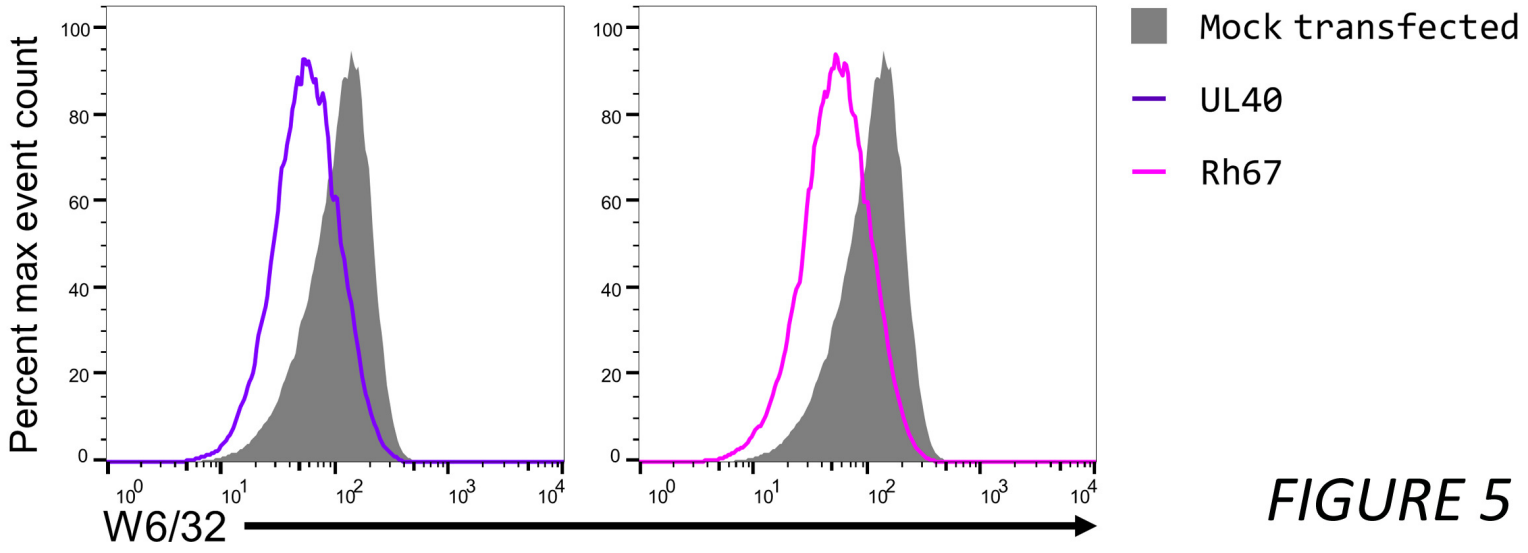
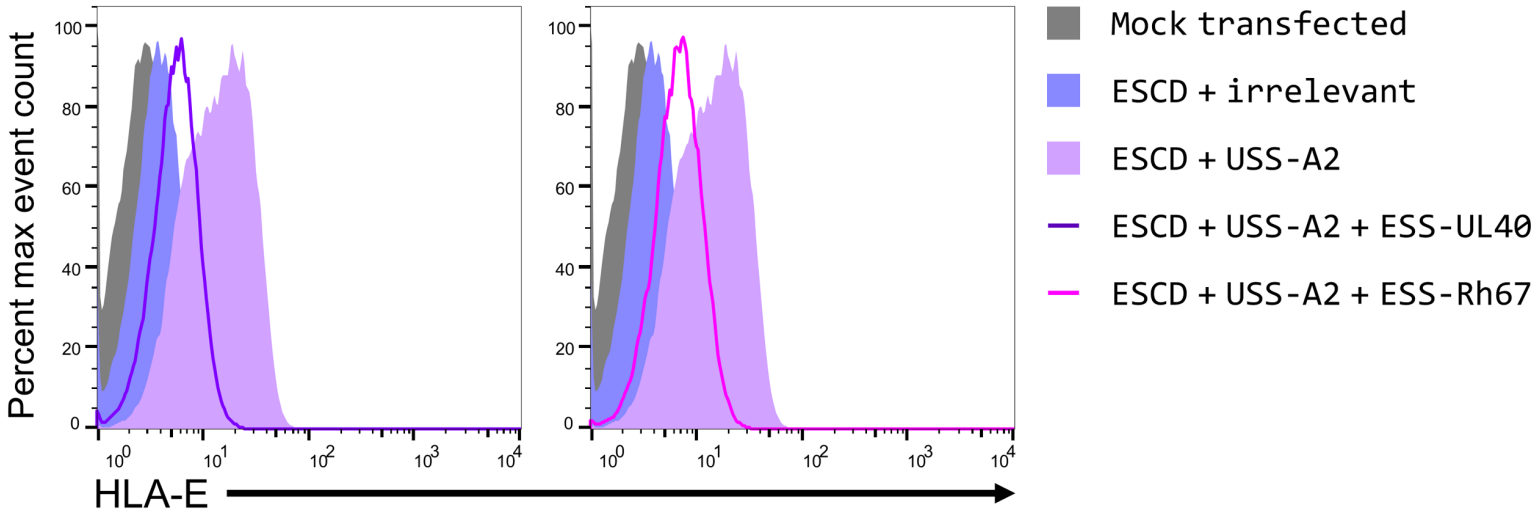

Percent max event count
W6/32
Mock transfected
UL40
Rh67

FIGURE 5

A


Percent max event count
100
80
60
40
20
0
10^0
10^1
10^2
10^3
10^4
HLA-E
Mock transfected
ESCD + irrelevant
ESCD + USS-A2
ESCD + USS-A2 + ESS-UL40
ESCD + USS-A2 + ESS-Rh67

B

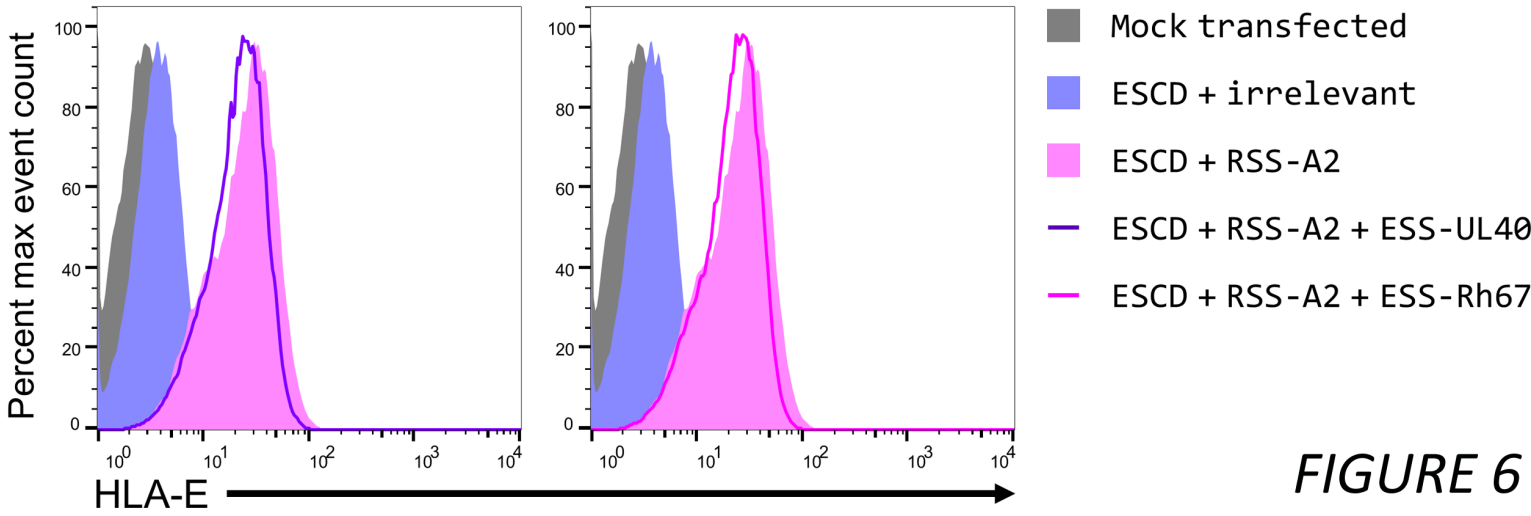

Percent max event count
100
80
60
40
20
0
10^0
10^1
10^2
10^3
10^4
HLA-E
Mock transfected
ESCD + irrelevant
ESCD + RSS-A2
ESCD + RSS-A2 + ESS-UL40
ESCD + RSS-A2 + ESS-Rh67

FIGURE 6