

1 **Filoviruses Require Endosomal Cysteine Proteases for Entry But Exhibit**  
2 **Distinct Protease Preferences**

3 **John Misasi<sup>1,2¶</sup>, Kartik Chandran<sup>1§\*</sup>, Jin-Yi Yang<sup>1</sup>, Bryden Considine<sup>1,2</sup>, Claire Marie**  
4 **Filone<sup>1,3</sup>, Marceline Côté<sup>1</sup>, Nancy Sullivan<sup>4</sup>, Giulia Fabozzi<sup>4</sup>, Lisa Hensley<sup>3</sup>, James**  
5 **Cunningham<sup>1,5\*</sup>**

6 **Running Title: Filoviruses require endosomal cysteine proteases**

7 <sup>1</sup>**Division of Hematology, Department of Medicine, Brigham & Women's Hospital, Boston,**  
8 **MA USA**

9 <sup>2</sup>**Division of Infectious Diseases, Children's Hospital Boston, Boston, MA USA**

10 <sup>3</sup>**U.S. Army Medical Research Institute of Infectious Diseases, Ft. Detrick, MD USA**

11 <sup>4</sup>**Vaccine Research Center, National Institute for Allergy and Infectious Diseases, National**  
12 **Institutes of Health, Bethesda, MD USA**

13 <sup>5</sup>**Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA**  
14 **USA**

15 <sup>§</sup>**Current affiliation: Department of Microbiology and Immunology, Albert Einstein College**  
16 **of Medicine, Bronx, NY 10461**

17 <sup>¶</sup>**These authors made equivalent contributions**

18 **\*Corresponding Authors:**

19 **James Cunningham:**

20 **Division of Hematology**

21 **Brigham and Women's Hospital**

22 **Karp Family Research Building, Rm 5213**

23 **1 Blackfan Circle**

24 **Boston, MA 02115**

25 **(617) 355-9058 (p) (617) 355-9124 (f) [jcunningham@rics.bwh.harvard.edu](mailto:jcunningham@rics.bwh.harvard.edu)**

26

27 **Kartik Chandran:**

28 **Department of Microbiology and Immunology**

29 **Albert Einstein College of Medicine**

30 **Forchheimer Building, Room 414**

31 **1300 Morris Park Ave**

32 **Bronx, NY 10461**

33 **(718) 430-8851 (p) (718) 430-8850 (f) [kartik.chandran@einstein.yu.edu](mailto:kartik.chandran@einstein.yu.edu)**

34

35 **Abstract Word Count: 226**

36 **Text Word Count: 3,223**

37

38 **ABSTRACT**

39 Filoviruses are enveloped viruses that cause sporadic outbreaks of severe hemorrhagic fever (6, 8,  
40 9, 16). Previous studies revealed that endosomal cysteine proteases are host factors for ebolavirus  
41 Zaire (7, 35). In this report, we show that infection mediated by glycoproteins from other  
42 phylogenetically-diverse filoviruses are also dependent on these proteases and provide additional  
43 evidence indicating they cleave GP1 and expose the binding domain for the critical host factor  
44 Niemann-Pick C1. Using selective inhibitors and knockout-derived cell lines, we show that the  
45 ebolaviruses Zaire and Cote d'Ivoire are strongly dependent on cathepsin B, while the ebolaviruses  
46 Sudan and Reston, and Marburg virus are not. Taking advantage of previous studies of cathepsin B  
47 inhibitor-resistant viruses (46), we found virus-specific differences in the requirement for cathepsin  
48 B are correlated with sequence polymorphisms at residues 47 in GP1 and 584 in GP2. We applied  
49 these findings to the analysis of additional ebolavirus isolates and correctly predicted that the  
50 newly-identified ebolavirus species Bundibugyo, containing D47 and I584, is cathepsin B-  
51 dependent and that ebolavirus Zaire-1995, the single known isolate of ebolavirus Zaire that lacks  
52 D47, is not. We also obtained evidence for virus-specific differences in the role for cathepsin L,  
53 including cooperation with cathepsin B. These studies strongly suggest that the use of endosomal  
54 cysteine proteases as host factors for entry is a general property of members of the family  
55 *Filoviridae*.

56

57 **INTRODUCTION**

58 Several lines of recent investigation have elucidated key steps in the pathway for entry of  
59 ebolaviruses into cells. Ebolavirus particles attach to cells through the binding of their  
60 glycoprotein (GP) to cell surface receptors or lectins, such as TIM-1 and DC-SIGN, expressed on  
61 the plasma membrane (1, 22, 27, 29, 37). Membrane-bound particles are taken up into cells by a  
62 macropinocytosis-like mechanism and transported to late endosomes/lysosomes (LE/LY) (20, 30,  
63 31, 34), which contain essential entry host factors. We previously showed that cleavage of  
64 ebolavirus Zaire-Mayinga (EBOV-May) GP by endosomal cysteine proteases is required for  
65 infection (7). More recent work has revealed a second host factor in LE/LY that is broadly required  
66 by filoviruses: Niemann-Pick C1 (NPC1) (5, 10), a multi-pass transmembrane protein that resides  
67 in the limiting membrane (44). According to a recently proposed model, virus GP is cleaved by  
68 endosomal cysteine proteases and binds to NPC1 (10)

69 Several studies have examined the role of protease cleavage in more detail for the EBOV-May.  
70 They show that cathepsin L functions in concert with cathepsin B to cleave the GP1 subunit of  
71 virus GP (7, 35). Structural and functional studies reveal that proteases remove the heavily  
72 glycosylated carboxyl-terminal domain of GP1 to expose a more conserved domain that is closely  
73 associated with GP2 (12, 19, 25) and that is proposed to contain the binding site for the filovirus  
74 receptor (4, 13, 24, 28). Further, we recently showed that cleaved, but not uncleaved GP1, binds to  
75 purified LE/LY membranes in an NPC1-dependent manner and co-immunoprecipitates with NPC1  
76 (10). We identified small molecules that target NPC1 and inhibit infection, and showed they block  
77 binding of cleaved GP1 (10), strongly suggesting the conserved N-terminal domain of GP1 is a  
78 ligand for NPC1. Taken together, these previous findings suggest a model in which proteolytic  
79 cleavage of GP to remove the carboxyl-terminal domain of GP1 and expose its N-terminal domain

80 may be functionally analogous to the role of CD4 binding to HIV gp120 to displace highly  
81 variable loops and create the co-receptor binding site (18). Our recent studies show that NPC1  
82 expression is also essential for infection by virus isolates from the other species that make up the  
83 family *Filoviridae* (23) including Cote d'Ivoire, Sudan, Reston, Bundibugyo and Marburg (5, 10).  
84 In this report, we have tested this model of infection by examining the endosomal cysteine  
85 protease requirements for infection by these viruses.

86

87 **MATERIALS AND METHODS**

88 **Cells and cell culture conditions.**

89 Vero, 293T and Mouse Embryonic Fibroblast (MEF) cell lines were maintained in Dulbecco's  
90 Modified Eagle's Media (DMEM), 100 µg/mL penicillin/streptomycin and 2mM L-glutamine  
91 (Invitrogen). 293T and Vero Cells were carried in either 10% FBS or 5% FBS + 5% FetalPlex  
92 (Gemini). Mouse embryonic fibroblasts (MEFs) were carried in 10% FBS. During virus  
93 production and at least 18 hours prior to infection Vero, 293T and MEF cells were grown in 10%  
94 FBS containing media. MEFs were described previously (7). CHO NPC1<sup>-/-</sup> and CHO NPC1<sup>-/-</sup>  
95 stably expressing mouse NPC1 were described previously (10).

96 **Expression plasmids.**

97 Expression plasmids for mouse cathepsin B, mouse cathepsin L, ebolavirus Zaire-Mayinga  
98 (EBOV-May) GP and mucin deleted ( $\Delta$ Muc) EBOV-May GP were described previously (7).  
99 Plasmids encoding ebolavirus Cote d'Ivoire-Ivory Coast GP, ebolavirus Sudan-Boniface GP,  
100 ebolavirus Reston-Pennsylvania GP and Lake Victoria marburgvirus-Musoke GP were obtained  
101 from Anthony Sanchez (CDC) and subcloned into pCAGGS.  $\Delta$ Muc GP constructs were created  
102 using PCR to remove the following amino acids: EBOV-May  $\Delta$ Muc GP ( $\Delta$ a.a. 309-489), Cote  
103 d'Ivoire $\Delta$ Muc GP ( $\Delta$ a.a. 310-489), Sudan $\Delta$ Muc GP ( $\Delta$ a.a. 309-490), Reston $\Delta$ Muc GP ( $\Delta$ a.a. 310-  
104 490), Bundibugyo $\Delta$ Muc GP ( $\Delta$ a.a. 309-489). This PCR added a silent Xba I site at the site of the  
105 deletion in each GP. Empty pCAGGS plasmid or plasmid encoding  $\beta$ -galactosidase were used as  
106 sham vector controls.

107 Chimeric EBOV-May GP1/Reston GP2 and Reston GP1/ EBOV-May GP2 were created by  
 108 restriction digest of the parental plasmids with Xba I (New England Biolabs) to liberate GP1 and  
 109 GP2 containing fragments for each virus GP. The fragments were gel purified and ligated using T4  
 110 DNA ligase (New England Biolabs). Vectors were verified by DNA sequencing.

111 EBOV-May  $\Delta$ Muc and Reston $\Delta$ Muc mutations were made using site directed PCR mutagenesis.  
 112 Primers sequences were: 5'-CTCTAGAGCCTCTGCTAACCA,  
 113 5'CATTACAGGTTAGTGAAGTCGACAACTAGTTTGT, 5'-  
 114 GTCGACTTCACTAACCTGTAATGT, 5'-ACTCTCAAAGCAACAGATATTGATCAATTGGT,  
 115 5'-TCAATATCTGTTGCTTTGAGAGT, 5'-  
 116 GCTACGCACCTTTTCACTCCTCAACCGTAAGGCAATTGA, 5'-  
 117 AGTGAAAAGGTGCGTAGCTCA, 5'-CTGAGGACTTACTCAATTCTTAACAGAAAAGCT,  
 118 5'-GAATTGAGTAAGTCCTCAGT, 5'-TATACTCGAGCTCGGTACCTCAAACCT, and 5'-  
 119 GCTAGCTCGAGAAATCAACACA. Chimeric viruses were made as described above.

120 **Infection assays.**

121 VSV pseudotyped viruses were created as previously described (7). Vesicular stomatitis virus  
 122 pseudotyped particles encoding GFP (VSV<sub>GFP</sub>) were added to cells in serial ten-fold dilutions and  
 123 assayed using fluorescent microscopy or flow cytometry. When using fluorescent microscopy,  
 124 GFP positive cells were counted manually and an infectious unit (i.u.) was defined as one GFP  
 125 expressing cell when within the linear range of dilutions. Titer is defined as i.u./ml of virus added  
 126 and is determined using the following formula: (# GFP expressing cells X dilution factor ) /  
 127 (volume (in mL) of virus added to the well). Two color flow cytometry was used to determine the  
 128 relative ratio of infected to uninfected MEF cathepsin B<sup>-/-</sup> L<sup>+/+</sup> cells in cells that were co-

129 transfected with mRFP expressing plasmid and a sham plasmid, mouse cathepsin B or mouse  
130 cathepsin L. Details of this protocol were described previously (7).

131 MEF cathepsin B<sup>-/-</sup> L<sup>-/-</sup> in rescue assays were performed in 6-well dishes. Sixteen to twenty hours  
132 after seeding the dishes, the media was changed to low serum (0.2%). After four hours, cells were  
133 transfected using 4 µg total DNA and 10 µL Lipofectamine 2000 (Invitrogen) that were diluted  
134 into OptiMEM (Invitrogen). The plasmids used were pCAGGS (4 µg), mouse cathepsin B (4 µg),  
135 mouse cathepsin L (4 µg), or both cathepsin B (2 µg) and cathepsin L (2 µg). Five hours following  
136 transfection, cells were washed once with PBS and fresh media containing 10% FBS added.  
137 Twenty-four hours post-transfection, cells were removed with Trypsin/EDTA solution  
138 (Invitrogen), spun down, resuspended in fresh media and eight-thousand cells added to each well  
139 of a 48-well dish. Following an additional 18 hours, the media was changed. Four hours later,  
140 pseudotyped VSV viruses encoding luciferase (VSV<sub>luc</sub>) were added at dilutions within the linear  
141 range of infection and sufficient to give relative luminescent units (RLU) signals between 10<sup>4</sup>-10<sup>5</sup>  
142 RLU on cells transfected with the pCAGGS. Twenty-four hours after the addition of the virus,  
143 cells were lysed using standard conditions for the firefly luciferase kit (Promega). Lysates were  
144 transferred to a 96-well plate, luciferin reagent added and after a 1 minute incubation RLU  
145 measured on an EnVison plate reader (Perkin Elmer).

146 **Protease inhibitors and protease activity assays.** Protease inhibitors E-64, E-64d and CA074  
147 (Sigma) were dissolved in dimethylsulfoxide (DMSO) and dispensed into culture medium  
148 immediately before use. The final DMSO concentration (v/v) in media was always 1%. Cell  
149 monolayers were preincubated with inhibitors or DMSO for 3 to 4 hours at 37°C. Viruses were  
150 added directly to the culture medium containing DMSO/inhibitor and infectivities were measured  
151 as described.

152 The enzymatic activities of cathepsin B and cathepsin L in acidified lysates (100mM NaCl, 50 mM  
153 NaAcetate, 0.5% Triton X-100, pH 5.5) of Vero cells and MEFs were assayed as previously  
154 described (7, 46). The absence of cathepsin B activity in MEF cathepsin B<sup>-/-</sup> cathepsin L<sup>+/+</sup>, the  
155 lack of cathepsin B and cathepsin L activity in MEF cathepsin B<sup>-/-</sup> cathepsin L<sup>-/-</sup> and the presence  
156 of similar levels of cathepsin L activity in cathepsin B<sup>-/-</sup> cathepsin L<sup>+/+</sup> and wild-type MEFs was  
157 confirmed.

158 **Ebolavirus Sudan infection.** Vero cells were treated with E-64d (300 μM) or vehicle (1%  
159 DMSO) for 4 hours, and then infected with ebolavirus Zaire-Mayinga or Sudan-Gulu  
160 (MOI=0.1). After 1 hour, the virus inoculums were removed by washing and fresh media  
161 containing E-64d or vehicle added. Cell supernatant was collected on day 3. RNA was isolated  
162 from the supernatant using Virus RNA Extraction kits (Qiagen) and ebolavirus NP RNA was  
163 measured using a quantitative RT-PCR assay (45). Virus titer was calculated using a standard  
164 curve obtained using a virus stock of known titer as determined by plaque assay.

165 **Marburg virus infection.** Vero cells were seeded into 6-well plates at a density of 2x10<sup>5</sup>  
166 cells/well. One day later, the media was removed and replace with fresh medium containing  
167 E64d (300 μM). After 4 hours, Ebola Zaire (MOI = 0.2) or Marburg virus Ci67 (Popp, MOI =  
168 0.2) was added to cells/drug. One hour later, medium containing drug/virus was removed, cells  
169 were washed with fresh medium, and fresh medium containing drug was added to each well.  
170 Time zero samples were harvested immediately, and remaining samples were incubated for 72  
171 hours. Supernatants were harvested for determination of virus by quantitative RT-PCR, and cells  
172 were lysed for determination of viral protein by western blot (14).

173 **Protease assay.** VSV virus particles bearing mucin domain deleted GPs from the indicated  
174 viruses were incubated in the presence of 0.2 mg/ml of chymotrypsin in reaction buffer (10 mM

175 Tris-HCl, 135 mM NaCl, 1 mM EDTA, pH 7.5). After 1 hour, the digestion was stopped by the  
176 addition of phenylmethanesulfonyl fluoride (PMSF) to a final concentration of 1 mM. Virus  
177 particles were deglycosylated by overnight incubation in the presence of PNGaseF (New  
178 England Biolabs). GP1 digestion products were analyzed under denaturing conditions via  
179 immunoblot against a conserved N-terminal region of ebolavirus GP1 using polyclonal rabbit  
180 anti-Ebola GP1 antibody (10).

181 **CHO cell infections.** VSV<sub>luc</sub> virus particles bearing GPs from the indicated ebolaviruses were  
182 digested in the presence or absence of chymotrypsin (as described above, but not  
183 deglycosylated). CHO NPC1<sup>-/-</sup> cells and CHO NPC1<sup>-/-</sup> cells stably expressing mouse NPC1 were  
184 incubated in the presence or absence of E-64d (300 μM). After four hours, virus particles were  
185 added to the cells. Sixteen hours after the addition of virus, luciferase activity was measured as  
186 described above.

187 **NPC1 membrane binding assay.** Expression and purification of EBOV-May<sub>ΔTM</sub> has been  
188 described previously (10). An expression vector encoding SUDV GP<sub>ΔTM</sub> that is fused to GCN4  
189 trimerization/His tag was also prepared (residues 1-309,491-657). Chymotrypsin cleaved SUDV  
190 GP<sub>ΔTM</sub> was created by incubation in the presence of 0.2 mg/ml chymotrypsin in chymotrypsin  
191 reaction buffer. After 30 minutes, the reaction was stopped by the addition of 1 mM EDTA, 1  
192 mM PMSF and 1 X EDTA-free complete protease inhibitor cocktail (Roche). Thermolysin  
193 digestion of EBOV-May<sub>ΔTM</sub> GP was described previously (10). Late endosomes/lysosomes  
194 (LE/LY) were isolated by differential centrifugation and Percoll (Sigma) density gradient  
195 centrifugation. LE/LY were disrupted by incubation with methionine methyl ester (Sigma) and  
196 coated onto high binding ELISA plates (Corning). Following attachment, unbound LE/LY  
197 membranes were removed and plates blocked with PBS + 5% FBS. Bound membranes were

198 incubated with the indicated amounts of native or protease-cleaved trimer in PBS + 5% FBS.  
 199 Unbound GP<sub>ΔTM</sub> protein was removed, membranes washed, and membrane bound GP<sub>ΔTM</sub> protein  
 200 recovered in SDS loading buffer and analyzed by immunoblot using GPI antiserum as described  
 201 previously (10).

202 **Cloning of Bundibugyo GP.** Trizol reagent (Invitrogen) inactivated ebolavirus Bundibugyo viral  
 203 RNA (vRNA) was precipitated using the standard protocol in the Trizol package insert. First strand  
 204 cDNA was created from this vRNA using random primers and SuperScript II (Invitrogen). Frame  
 205 shifted Bundibugyo GP was amplified via PCR using Phusion HS (New England Biolabs) and the  
 206 following primers: 5'- TAAATGCATGGTTACATCAGGA, 5'-  
 207 TGTGAAGTTCTTCTTATTTTCCCAGAAGGC, 5'- TAAGAAGAACTTCACAAAACCCCT  
 208 and 5'- TATCTCGAGGACTAGATTAGAGTAGA. The final PCR product was cloned into  
 209 pCAGGS MCS using Nsi I and Xho I. The clone was verified by sequencing. A mucin deleted  
 210 version of this plasmid was created by PCR with the previous primers, 5'-  
 211 TCTCCGACATATGGTACCGCAAATCTGCTGACAGGCTCA and 5'-  
 212 GGTACCATATGTCGGAGAGGTACCGACAGACAGCTCTTCA, and cloning into pCAGGS  
 213 MCS with Nsi I and Xho I. This plasmid was restricted with Kpn I and religated to create the  
 214 mucin deleted (Δ a.a. 309-489) BDBV. Plasmids were verified by sequencing. All restriction  
 215 enzymes were obtained from New England Biolabs.

216

217 **RESULTS**

218 **Endosomal cysteine proteases are host factors for filovirus entry.** Previous studies show that  
219 EBOV-May infection is reduced by >99% by E-64, a highly specific small molecule inhibitor of  
220 endosomal cysteine proteases including cathepsin L and cathepsin B (7, 35). We used E-64 as a  
221 probe to analyze the host requirements for entry by GPs from other filoviruses. Vero cells were  
222 pre-treated with E-64 and then challenged with VSV vectors pseudotyped with GPs from EBOV-  
223 May, ebolavirus Cote d'Ivoire-Ivory Coast (CIEBOV), ebolavirus Sudan-Boniface (SUDV),  
224 ebolavirus Reston-Pennsylvania (RESTV) and Lake Victoria marburgvirus-Musoke (MARV).  
225 The titers of VSV particles pseudotyped with EBOV-May, CIEBOV, SUDV and MARV GP are  $6$   
226  $\times 10^9$ - $1 \times 10^{10}$  i.u./mL and the titer of VSV RESTV GP particles is ten-fold lower ( $1.3 \times 10^8$   
227 i.u./mL) (Fig. 1A). We found that E-64 treatment of target cells reduced infection by these viruses  
228 by >99.5% under conditions where infection by VSV G infection was reduced by <50% (Fig. 1A).  
229 These findings were confirmed in studies of viruses bearing GPs lacking the mucin-rich domain in  
230 GP1. To verify that cysteine proteases are bona fide host factors, we tested the effect of E-64d on  
231 growth of ebolavirus Zaire-Mayinga, ebolavirus Sudan-Gulu, and Lake Victoria marburgvirus-  
232 Musoke in Vero cells. The production of new virus was reduced by more than 99% in the presence  
233 of E-64d as determined by quantitative RT-PCR (Fig. 1B/C). Taken together, the results of these  
234 studies indicate that infection by EBOV-May, CIEBOV, SUDV, RESTV and MARV is dependent  
235 on endosomal cysteine proteases sensitive to E-64.

236 **Proteases target virus GP1.** Analysis of the amino acid sequences of GPs from EBOV-May,  
237 CIEBOV, SUDV, RESTV and MARV suggests the domain structures of GP1 and GP2 are  
238 conserved (data not shown), a conclusion that has been recently confirmed for Sudan GP (11, 25).  
239 Biochemical studies of EBOV-May indicate that endosomal cysteine proteases remove the heavily

240 glycosylated carboxyl-terminal "cap" domain and expose the N-terminal domain (7, 12, 19, 24, 25,  
241 35) that is the ligand for the essential host factor NPC1 (10). To determine if a protease-sensitive  
242 carboxyl-terminal domain in GP1 is conserved among ebolaviruses, we incubated VSV particles  
243 pseudotyped with GPs from EBOV-May, CIEBOV, SUDV and RESTV with chymotrypsin and  
244 analyzed the effect on virus infection. Chymotrypsin is a serine endoprotease that faithfully  
245 mimics the action of cathepsin L (46). We observed that like EBOV-May, chymotrypsin digestion  
246 of CIEBOV, SUDV and RESTV removes the carboxyl-terminal domains of GP1 to create an N-  
247 terminal 18-20 kDa fragment (Fig. 2A). We analyzed the function of chymotrypsin-cleaved virus  
248 particles on Chinese hamster ovary (CHO) cells and found that they are infectious and dependent  
249 on both NPC1 and cysteine protease activity (Fig. 2B). We next analyzed binding of Sudan GP to  
250 NPC1 membranes. The source of GP is an ectodomain trimer in which the transmembrane domain  
251 is replaced with a GCN4 trimerization domain (SUDV<sub>ΔTM</sub>). We incubated uncleaved and  
252 chymotrypsin cleaved SUDV<sub>ΔTM</sub> GP with purified late endosome/lysosome membranes (LE/LY)  
253 from knockout and NPC1 expressing CHO cells. We found that, like EBOV-May<sub>ΔTM</sub> GP,  
254 SUDV<sub>ΔTM</sub> GP binding to membranes is dependent on cleavage of GP1 and expression of NPC1  
255 (Fig. 2C). Although more work is needed, these findings suggest that at least one function of  
256 endosomal cysteine proteases in filovirus infection is to remove the carboxyl-terminal domain of  
257 GP1 and expose the NPC1 binding site.

258 **Requirement for cathepsin B is not conserved.** EBOV-May infection is strongly dependent on  
259 the endosomal cysteine protease, cathepsin B (7, 35). To test the hypothesis that cathepsin B  
260 activity is required for infection by other filoviruses, we measured VSV GP-dependent infection of  
261 Vero cells treated with the selective cathepsin B inhibitor, CA074. We confirmed that EBOV-May  
262 GP-dependent infection is reduced as a function of the concentration of CA074 (0-80 μM) and is

263 closely correlated with cathepsin B activity (Fig. 3A/B). We observed that like EBOV-May,  
264 CIEBOV GP-mediated infection is also closely correlated with cathepsin B activity. However,  
265 SUDV GP infection was not as sensitive to CA074 inhibition as EBOV-May or CIEBOV GP and  
266 neither RESTV nor MARV GP-mediated infection was significantly reduced when cathepsin B  
267 activity was inhibited (Fig. 3A/B).

268 As an independent test of the role of cathepsin B in infection, we studied murine fibroblast cells  
269 (MEFs) derived from cathepsin B knockout mice (cathepsin B<sup>-/-</sup>). The pattern of infection of  
270 cathepsin B<sup>-/-</sup> MEFs by filovirus GPs is closely correlated with the pattern of infection of Vero  
271 cells treated with CA074: EBOV-May and CIEBOV GP infection of cathepsin B<sup>-/-</sup> MEFs is <1%  
272 of wild type MEFs, SUDV (7%), RESTV (25-40%) and MARV infection (65-99%) (Fig. 3C/D).  
273 The introduction of an expression plasmid encoding cathepsin B into the cathepsin B<sup>-/-</sup> MEFs  
274 enhanced infection by each of the viruses to 90-150% of infection of wild type MEFs (Fig. 3D)  
275 thus confirming that the defect in infection is due to cathepsin B deficiency. These findings  
276 demonstrate that dependence on cathepsin B by EBOV-May and CIEBOV is not a conserved  
277 property of all filoviruses.

278 **Cathepsin L.** Cathepsin L is an E-64-sensitive endopeptidase that is expressed in cells expressing  
279 cathepsin B (15, 17, 32, 40, 42, 43). To investigate the role of cathepsin L in infection by  
280 cathepsin B dependent and independent viruses, we studied infection of Vero cells treated with the  
281 inhibitor FYdmk. At low concentrations where cathepsin L, but not Cat B, activity is blocked (<1  
282 μM), there was no effect of FYdmk on infection of ebolavirus or marburgvirus pseudotyped  
283 particles (Fig. 4A/B). However, when the concentration of FYdmk exceeded >1 μM and cathepsin  
284 B and likely other endosomal cysteine proteases are also inhibited, infection by each virus was  
285 reduced, consistent with the results of studies using the general cysteine protease inhibitor, E-64.

286 These results indicate that cathepsin L is not essential for either the cathepsin B dependent or  
287 independent viruses.

288 In cathepsin B<sup>-/-</sup> MEFs, overexpression of cathepsin L enhanced infection by SUDV, RESTV  
289 and MARV but was unable to overcome the defect imposed by loss of cathepsin B activity for  
290 EBOV-May or CIEBOV (Fig. 3D). To determine if cathepsin L is able to support filovirus  
291 infection in the absence of cathepsin B, we used mouse fibroblasts obtained from cathepsin B<sup>-/-</sup>  
292 L<sup>-/-</sup> mice. In our initial experiments, we found that infection of these cells by CIEBOV, SUDV  
293 and RESTV closely corresponds to infection of cathepsin B<sup>-/-</sup> MEFs and CA074-treated Vero  
294 cells (Fig 4C). However, unlike cathepsin B<sup>-/-</sup> MEFs and Vero cells treated with low dose  
295 FYdmk, MARV infection of cathepsin B<sup>-/-</sup> L<sup>-/-</sup> cells was reduced by >95%. The transfection  
296 efficiency of cathepsin B<sup>-/-</sup> L<sup>-/-</sup> MEFs was too low to assess the role of cathepsin L and cathepsin  
297 B overexpression using VSV<sub>GFP</sub> particles. To circumvent this limitation, we utilized VSV<sub>luc</sub>  
298 pseudotyped particles at moi>1, which enhances the sensitivity of the measurement of virus  
299 infection. As expected, expression of cathepsin B was markedly superior to expression of  
300 cathepsin L in supporting infection by EBOV-May and CIEBOV (Cat B: 9-30-fold vs. Cat L: <  
301 3-fold). Remarkably, cathepsin B-dependent infection was markedly increased when cathepsin  
302 L was also expressed (Fig. 4D). These findings are consistent with the results of previous studies  
303 for EBOV-May (7) and indicate that expression of cathepsin L is not a substitute for the essential  
304 role of cathepsin B in EBOV-May and CIEBOV infection. In contrast, expression of either  
305 cathepsin L or cathepsin B enhanced RESTV infection of cathepsin B<sup>-/-</sup> L<sup>-/-</sup> MEFs. A consistent  
306 observation was that the response of SUDV to overexpression of cathepsin L and/or cathepsin B  
307 was not as robust as for the other viruses. In contrast to the studies of the ebolaviruses, cathepsin  
308 L (90-fold) was much more active than cathepsin B (2.5-fold) in supporting MARV infection of

309 these cells (Fig. 4D). Taken together, these findings identify marked virus-specific variation in  
310 the effect of cathepsin L expression on filovirus infection.

311 **Mapping GP determinants of cathepsin B.** EBOV-May and RESTV are closely related but  
312 differ in dependence on cathepsin B and therefore provided an opportunity to map the virus  
313 determinants of the cathepsin B requirement. To this end, we analyzed virus particles bearing  
314 chimeric GPs created by reciprocal exchange of portions of EBOV-May and RESTV GPs. In our  
315 initial experiment, the effect of exchanging GP1 and GP2 was examined. We observed that the  
316 titer of EBOV-GP1/RESTV-GP2 particles on Vero cells was similar to RESTV and the titer of  
317 RESTV-GP1/EBOV-GP2 particles was similar to EBOV-May, indicating that the chimeric GPs  
318 were functional (Fig. 5A). We observed that neither of the virus particles expressing chimeric  
319 glycoproteins EBOV/RESTV or RESTV/EBOV were as sensitive as EBOV-May particles to  
320 inhibition by CA074 as to E-64. These findings indicate that the determinants of cathepsin B  
321 dependence are not localized exclusively within either GP1 or GP2. A previous study of EBOV-  
322 May-derived viruses selected for resistance to CA074 suggested an explanation. This study  
323 showed that single amino acid changes in residues clustered either near the N-terminus of GP1  
324 (N40K/S/T, T42A, L43F or D47V) or in the hr1 segment of GP2 (I584F or K588R) conferred  
325 resistance to CA074 (46). Using these findings as a guide, we noted that RESTV GP differs from  
326 EBOV-May GP at residues 47 (D→E) in GP1 and 584 (I→L) in GP2 (Fig. 5B). To determine if  
327 these subtle changes mediate the difference in behavior of these viruses, we exchanged these  
328 residues in EBOV-May and RESTV and tested the effect on sensitivity to CA074. We found that  
329 substitution of glutamic acid for D47 or leucine for I584 alone or in combination conferred  
330 resistance of EBOV-May to CA074. However, the introduction of the reciprocal changes E48D  
331 and/or L585I into RESTV GP did not confer a requirement for cathepsin B (Fig. 5C/D). These

332 findings indicate that D47 and I584 are necessary for EBOV-May but not sufficient for RESTV to  
333 depend on cathepsin B as a host factor for infection of Vero cells.

334 **Sequence analysis predicts cathepsin B dependence.** The discovery that cathepsin B dependence  
335 of EBOV-May maps to residues D47 and I584 suggested the possibility that these residues might  
336 also predict the protease requirements for other filoviruses. Indeed, we noted that D47 and I584 are  
337 conserved in EBOV-May and CIEBOV GP which are cathepsin B dependent, but not in SUDV  
338 (E47) or MARV (N47 and L584), which are not. A search of the NCBI database identified 27  
339 partial and complete sequences of GPs obtained from independent isolates classified as EBOV.  
340 Only one of the virus isolates, EBOV-1995 (Genbank ID-AY354458), differed in the cathepsin B-  
341 determining residues. Excluding the mucin regions, which are deleted from the GPs in this  
342 analysis, EBOV-May and EBOV-1995 differ only at residues 47 (D47E) and 544 (I544T), which  
343 had not been identified in previous studies of cathepsin B requirements (46). We prepared and  
344 analyzed VSV EBOV-May GP-derived particles containing E47 or T544 GP. As expected, these  
345 viruses are highly infectious and sensitive to E-64. However, inhibition of cathepsin B by CA074  
346 reduced the titer of T544 GP particles by >99% under conditions where the titer of E47 GP  
347 particles was minimally changed (Fig.6A). Thus, escape from dependence on cathepsin B was  
348 closely correlated with E47 and not T544. Over the time period since we began these  
349 investigations, an outbreak of hemorrhagic fever occurred in Uganda due to a filovirus that is  
350 sufficiently different from EBOV, CIEBOV, RESTV and MARV to be classified as the new  
351 species ebolavirus Bundibugyo (BDBV)(39). Analysis of the BDBV sequence revealed the  
352 presence of D47 and I584, thus predicting that BDBV is cathepsin B dependent. Indeed, we found  
353 that infection by VSV BDBV GP particles was inhibited by increasing concentrations of CA074  
354 (0-80  $\mu$ M) in a pattern closely correlated with EBOV-May (Fig. 6B). In addition, we found that

355 infection by BDBV pseudotypes was blocked by treatment of cells with E-64 and in cathepsin B<sup>-/-</sup>  
356 MEFs (data not shown). Thus, in our limited cohort of filoviruses analyzed in this study, the  
357 presence of D47 correlated with dependence on cathepsin B for infection.

## 358 **DISCUSSION**

359 Over the past six years, significant progress has been made in understanding how filoviruses gain  
360 entry into cells and a model of infection based on this progress has been proposed (10). EBOV-  
361 May particles attach to lectins on the cell surface and are taken up by macropinocytosis into  
362 vesicles that are transported to LE/LY containing endosomal cysteine proteases and NPC1 (5, 10,  
363 20, 31, 32, 34, 43). One function of endosomal cysteine proteases is to cleave the carboxyl-  
364 terminal "cap" region of GP1 from the mushroom shaped GP protruding from the virus membrane  
365 and expose the stalk containing the protease-resistant N-terminal domain in GP1 that is the ligand  
366 for the Niemann-Pick C1 (NPC1) protein (4, 10-12, 19, 24, 25, 28). A second function may be to  
367 biochemically destabilize the GP protein, thereby sensitizing it to triggering for viral membrane  
368 fusion (3, 46). Further studies are needed to determine the roles of NPC1 (5, 10) and additional  
369 events including further cleavage of GP1 in this process (3, 11, 21, 35, 46).

370 In this report, we provide evidence that key aspects of this scheme are conserved among other  
371 filoviruses. The new findings show that like EBOV-May, the carboxyl-terminal domain of SUDV  
372 GP1 is a substrate for proteolytic cleavage, that cleaved particles are infectious and NPC1  
373 dependent, and that the protease-resistant N-terminal domain of GP1 binds to purified LE/LY  
374 membranes in an NPC1-dependent manner. These findings are consistent with the recent report  
375 showing that the domain organization of EBOV-May is conserved in SUDV (11). Moreover, we  
376 find that the sensitivity of the C-terminal domain of GP1 to cleavage and the dependence of

377 cleaved particles on NPC1 is conserved among other ebolaviruses. Consistent with this view, we  
378 find that isolates from each species of *Filoviridae* are E-64 sensitive, including growth of EBOV-  
379 May, SUDV and MARV. While more work needs to be done including studies of other proteases  
380 and MARV, the results of the experiments in this report, coupled with alignment of primary amino  
381 acid sequences of GPs which indicate that the domain structure is likely to be conserved (data not  
382 shown), suggest that cleavage and binding are key steps in the filovirus entry pathway. In this  
383 model, endosomal cysteine proteases are required for the efficient removal of the carboxyl-  
384 terminal domain of GP1 to expose the NPC1 binding domain. Endosomal cysteine proteases may  
385 mediate the additional steps necessary for orderly deployment of the virus membrane fusion  
386 activity (21, 35, 46). Indeed, two recent reports suggest that cysteine protease cleavage of the  $\beta$ -13-  
387 14 loop in GP1 promotes release of the GP2 fusion peptide (3,11). The studies in this report  
388 provide the basis for future studies to compare the virus requirements for specific proteases to  
389 cleave GP1, to bind to NPC1 and to release the GP2 fusion peptide.

390 The endosomal cysteine proteases are a family of 11 acid dependent proteases that reside in late  
391 endosome and lysosomes (15, 17, 32, 40-43). At present, little is known about the role individual  
392 members of this family play during filovirus infection. Previously we showed that EBOV-May  
393 required cathepsin B, but not cathepsin L, for entry (7). We confirm this finding and show that  
394 CIEBOV has the same requirements. Additional studies are needed to determine if the virus  
395 sensitivity to the absence of cathepsin B activity is due to a specific requirement for the double-  
396 chain isoform of cathepsin B (36). Although cathepsin L is not essential for any of the filoviruses  
397 studied, we found that it works in concert with cathepsin B to enhance infection of EBOV-May,  
398 CIEBOV and RESTV. Furthermore, cathepsin L activity is required for MARV infection of MEFs  
399 but not Vero cells, suggesting that the role of cathepsin L may be shared by other cysteine

400 proteases with endopeptidase activity. Remarkably, RESTV infection is sensitive to E-64 but is  
401 not sensitive to the loss of cathepsin B and L. This suggests that one or more additional E-64  
402 sensitive proteases can support RESTV infection.

403 The substrate specificity of endosomal cysteine proteases are largely governed by accessibility of  
404 the polypeptide chain to the active site of the protease and not by strong preference for specific  
405 sequence motifs (32, 41, 42). Thus, one consequence of the extensive variation in the sequence of  
406 the carboxyl-terminal domain of GP1 is that it may alter the repertoire of cysteine proteases that  
407 are able to cleave GP1. Therefore, the presence of multiple proteases with overlapping substrate  
408 preferences in late endosomes and lysosomes of host cells may provide for redundancy in the  
409 conditions for cleavage of GP and might explain the virus-specific differences in dependence on  
410 cathepsin B and cathepsin L observed in our studies. One advantage of this scheme may be that  
411 effective cleavage of the GP1 cap and/or fusion peptide release, is maintained in presence of  
412 selective pressure from host immune recognition for sequence diversification, analogous to the  
413 function of variable loops in HIV gp120 (2, 18). In this model, adaptation to loss of cathepsin B  
414 activity by a change to a single amino acid residue (i.e. D47, I584) may provide a means for rapid  
415 response to changes in endosomal cysteine protease expression between hosts or cell types.  
416 Sequence polymorphisms that specifically predict host factor preference have been identified in  
417 other virus envelope glycoproteins including SARS GP N479K interactions with human or civet  
418 ACE2 receptor, influenza HA1 interactions with  $\alpha$ -2 glycan linkage in human or  $\alpha$ -2-6 glycan  
419 linkage in avian sialic acid, and HIV gp120 binding to receptor CXCR4 and/or CCR5 (2, 26, 33,  
420 38). Analysis of the filoviruses identified in future outbreaks will provide further test of the utility  
421 of using the D47/I584 polymorphisms in GP in determining virus preference for host endosomal  
422 cysteine proteases.

423 **ACKNOWLEDGEMENTS**

424 This work was supported by grants from U54 AI057159, R01 CA104266 to JC, and PIDS-Sanofi-  
425 Pasteur Fellowship, K12-HD052896 and 5K08AI079381 to JM. KC was supported by a  
426 postdoctoral fellowship from the New England Research Center of Excellence in Biodefense and  
427 Emerging Infectious Diseases (NERCE/BEID). CF was supported by the Postgraduate Research  
428 Participation Program at the U.S. Army Medical Research and Material Command administered by  
429 the Oak Ridge Institute for Science and Education through an interagency agreement between the  
430 U.S. Department of Energy and USAMRMC. MC was supported by Fonds de la recherche en  
431 Santé du Québec (FRSQ).

432 We thank Scott Aoki, Anna Bruchez, Brenna Hill and Daniel Douek for assistance.

433 Opinions, interpretations, conclusions and recommendations are those of the authors and are not  
434 necessarily endorsed by the U.S. Department of Defense, the U.S. Department of the Army, or the  
435 U.S. Department of Health and Human Services.

436

437 **REFERENCES**

- 438 1. **Alvarez C. P., F. Lasala, J. Carrillo, O. Muñiz, A. L. Corbí, and R. Delgado.** 2002. C-type  
439 lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. *Journal*  
440 *of Virology* **76**:6841-4.
- 441 2. **Arthos J., K. C. Deen, M. A. Chaikin, J. A. Fornwald, G. Sathe, Q. J. Sattentau, P. R.**  
442 **Clapham, R. A. Weiss, J. S. McDougal, C. Pietropaolo, R. Axel, A. Trunch, P. J. Maddon,**  
443 **R.W. Sweet.** 1989. Identification of the residues in Human CD4 critical for binding HIV. *Cell*  
444 **57**:469-81.
- 445 3. **Brecher M., K. Schornberg, S. Delos, M. Fusco, E. Saphire, J. White.** 2011. Cathepsin  
446 Cleavage Potentiates the Ebola Virus Glycoprotein to Undergo a Subsequent Fusion Relevant  
447 Conformational Change. *Journal of Virology*. Oct 26. Epub ahead of print.
- 448 4. **Brindley M., L. Hughes, A. Ruiz, P. McCray, A. Sanchez, D. Sanders, W. Maury.** 2007.  
449 Ebola virus glycoprotein 1: Identification of residues important for binding and postbinding  
450 events. **81**: 7702-09.
- 451 5. **Carette J. E., M. Raaben, A. Wong, A. C. Herbert, G. Obernosterer, N. Mulherkar, A. I.**  
452 **Kuehne, P. J. Kranzusch, A. M. Griffin, G. Ruthel, P. Dal Cin, J. M. Dye, S. P. Whelan,**  
453 **and T. Chandran, K Brummelkamp.** 2011. Ebola virus entry requires the cholesterol  
454 transporter Niemann-Pick C1. *Nature* **477**: 340-3.
- 455 6. **Centers for Disease Control and Prevention.** 2001. Outbreak of Ebola hemorrhagic fever  
456 Uganda, August 200-January 2001. *MMWR* **50**:73-7.

- 457 7. **Chandran K., N. J. Sullivan, U. Felbor, S. P. Whelan, and J. M. Cunningham.** 2005.  
458 Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science* (New  
459 York, N.Y.) **308**:1643-5.
- 460 8. **Colebunders R. and M. Borchert.** 2000. Ebola haemorrhagic fever--a review. *J Infect* **40**:16-  
461 20.
- 462 9. **Colebunders R., A. Tshomba, M. D. Van Kerkhove, D. G. Bausch, P. Campbell, M.**  
463 **Libande, P. Pirard, F. Tshioko, S. Mardel, S. Mulangu, H. Sleurs, P. E. Rollin, J. J.**  
464 **Muyembe-Tamfum, B. Jeffs, M. Borchert.** 2007. Marburg hemorrhagic fever in Durba and  
465 Watsa, Democratic Republic of the Congo: clinical documentation, features of illness, and  
466 treatment. *J Infect Dis* **196** Suppl 2:S148-53.
- 467 10. **Côté M., J. Misasi, T. Ren, A. Bruchez, K. Lee, C. M. Filone, L. Hensley, Q. Li, D. Ory,**  
468 **K. Chandran and J. Cunningham.** 2011. Small molecule inhibitors reveal Niemann-Pick C1 is  
469 essential for ebolavirus infection. *Nature* **477**:34-8.
- 470 11. **Dias J., A. Kuehne, D. Abelson, S. Bale, A. Wong, P. Halfmann, M. Muhammad, M.**  
471 **Fusco, S. Zak, E. Kang, Y. Kawaoka, K. Chandran, J. Dye, E. Saphire.** 2011. A shared  
472 structural solution for neutralizing ebolaviruses. *Nature Structural & Molecular Biology* Nov 20.  
473 Advanced online publication.
- 474 12. **Dube D., M. B. Brecher, S. E. Delos, S. C. Rose, E. W. Park, K. L. Schornberg, J. H.**  
475 **Kuhn, and J. M. White.** 2009. The primed ebolavirus glycoprotein (19-kilodalton GP1,2):  
476 sequence and residues critical for host cell binding. *Journal of Virology* **83**:2883-91.

- 477 13. **Dube D., K. Schornberg, C. Shoemaker, S. Delos, T. Stantchev, K. Clouse, C. Broder, J.**  
478 **White.** 2010. Cell adhesion-dependent membrane trafficking of a binding partner for the  
479 ebolavirus glycoprotein is a determinant of viral entry. *Proceedings of the National Academy of*  
480 *Sciences* **107**: 16637-42.
- 481 14. **Fabozzi G., C. Nabel, M. Dolan, N. Sullivan.** 2011. Ebolavirus proteins suppress the effects  
482 of small interfering RNA by direct interaction with the mammalian RNA interference pathway. *J*  
483 *Virology*. **85**:2512-23.
- 484 15. **Fujishima A., Y. Imai, T. Nomura, Y. Fujisawa, Y. Yamamoto, and T. Sugawara.** 1997.  
485 The crystal structure of human cathepsin L complexed with E-64. *FEBS Letters* **407**:47-50.
- 486 16. **Geisbert T. W., and P. B. Jahrling.** 2004. Exotic emerging viral diseases: progress and  
487 challenges. *Nature Medicine* **10**:S110-21.
- 488 17. **Guncar G., G. Pungercic, I. Klemencic, V. Turk, and D. Turk.** 1999. Crystal structure of  
489 MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for  
490 differentiation between cathepsins L and S. *The EMBO Journal* **18**:793-803.
- 491 18. **Harrison S. C.** 2008. Viral membrane fusion. *Nature Structural & Molecular Biology*  
492 **15**:690-8.
- 493 19. **Hood C. L., J. Abraham, J. C. Boyington, K. Leung, P. D. Kwong, and G. J. Nabel.**  
494 2010. Biochemical and structural characterization of cathepsin L-processed Ebola virus  
495 glycoprotein: implications for viral entry and immunogenicity. *Journal of Virology* **84**:2972-82.

- 496 20. **Hunt C. L., A. a Kolokoltsov, R. a Davey, and W. Maury.** 2011. The Tyro3 receptor  
 497 kinase Axl enhances macropinocytosis of Zaire ebolavirus. *Journal of Virology* **85**:334-47.
- 498 21. **Kaletsky R., G. Simmons, P. Bates.** 2007. Proteolysis of the ebola virus glycoproteins  
 499 enhances virus binding and infectivity. *Journal of Virology* **81**:13378-84.
- 500 22. **Kondratowicz A. S., N. J. Lennemann, P. L. Sinn, R. A. Davey, C. L. Hunt, S. Moller-**  
 501 **Tank, D. K. Meyerholz, P. Rennert, R. F. Mullins, M. Brindley, L. M. Sandersfeld, K.**  
 502 **Quinn, M. Weller, P. B. McCray, J. Chiorini, and W. Maury.** 2011. T-cell immunoglobulin  
 503 and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus.  
 504 *Proceedings of the National Academy of Sciences of the United States of America* **108**:8426-31.
- 505 23. **Kuhn J. H., S. Becker, H. Ebihara, T. W. Geisbert, K. M. Johnson, T. Kawaoka, W. I.**  
 506 **Lipkin, A. I. Negredo, S. V. Netesov, S. T. Nichol, G. Palacios, C. J. Peters, A. Tenorio, V.**  
 507 **E. Volchkov, P. B. Jahrling.** 2010. Proposal for a revised taxonomy of the family *Filoviridae*:  
 508 classification, names of taxa and viruses, and virus abbreviations. *Archives of Virology*  
 509 **155**:2083-103.
- 510 24. **Kuhn J. H., S. R. Radoshitzky, A. C. Guth, K. L. Warfield, W. Li, M. J. Vincent, J. S.**  
 511 **Towner, S. T. Nichol, S. Bavari, H. Choe, M. J. Aman, and M. Farzan.** 2006. Conserved  
 512 receptor-binding domains of Lake Victoria marburgvirus and Zaire ebolavirus bind a common  
 513 receptor. *The Journal of Biological Chemistry* **281**:15951-8.
- 514 25. **Lee J. E., M. L. Fusco, A. J. Hessel, W. B. Oswald, D. R. Burton, and E. O. Saphire.**  
 515 2008. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor.  
 516 *Nature* **454**:177-82.

- 517 26. **Li F., W. Li, M. Farzan, S. C. Harrison.** 2005. Structure of SARS Coronavirus spike  
 518 receptor-binding domain complexed with receptor. *Science* **309**: 1864-68.
- 519 27. **Lin G., G. Simmons, S. Pöhlmann, F. Baribaud, H. Ni, G. J. Leslie, B. S. Haggarty, P.**  
 520 **Bates, D. Weissman, J. A. Hoxie, and R. W. Doms.** 2003. Differential N-linked glycosylation  
 521 of human immunodeficiency virus and Ebola virus envelope glycoproteins modulates  
 522 interactions with DC-SIGN and DC-SIGNR. *Journal of Virology* **77**:1337-46.
- 523 28. **Manicassamy B., J. Wang, H. Jiang, L. Rong.** 2005. Comprehensive Analysis of ebola  
 524 virus GP1 in viral entry. *Journal of Virology*. **79**: 4793-4805.
- 525 29. **Marzi A., A. Akhavan, G. Simmons, T. Gramberg, H. Hofmann, P. Bates, V. R.**  
 526 **Lingappa, and S. Pöhlmann.** 2006. The signal peptide of the ebolavirus glycoprotein influences  
 527 interaction with the cellular lectins DC-SIGN and DC-SIGNR. *Journal of Virology* **80**:6305-17.
- 528 30. **Mulherkar M. M. Raaben, J.C. de la Torre, S. Whelan, K. Chandran.** 2011. The ebola  
 529 virus glycoprotein mediates entry via a non-classical dynamin-dependent macropinocytic  
 530 pathway. *Virology* **416**: 72-83.
- 531 31. **Nanbo A., M. Imai, S. Watanabe, T. Noda, K. Takahashi, G. Neumann, P. Halfmann,**  
 532 **and Y. Kawaoka.** 2010. Ebolavirus is internalized into host cells via macropinocytosis in a viral  
 533 glycoprotein-dependent manner. *PLoS Pathogens* **6**.
- 534 32. **Otto H.-H., and T. Schirmeister.** 1997. Cysteine Proteases and Their Inhibitors. *Chemical*  
 535 *Reviews* **97**:133-172.

- 536 33. **Ribeiro R. M., M. D. Hazenberg, A. S. Perelson, and M. P. Davenport.** 2006. Naïve and  
537 memory cell turnover as drivers of CCR5-to-CXCR4 tropism switch in human  
538 immunodeficiency virus type 1: implications for therapy. *Journal of Virology* **80**:802-9.
- 539 34. **Saeed M. F., A. Kolokoltsov, T. Albrecht, and R. Davey.** 2010. Cellular Entry of Ebola  
540 Virus Involves Uptake by a Macropinocytosis-Like Mechanism and Subsequent Trafficking  
541 through Early and Late Endosomes. *PLoS Pathogens* **6**:e1001110.
- 542 35. **Schorner K., S. Matsuyama, K. Kabsch, S. Delos, A. Bouton, and J. White.** 2006. Role  
543 of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *Journal of Virology*  
544 **80**:4174-8.
- 545 36. **Schorner K., C. Shoemaker, D. Dube, M. Abshire, S. Delos, A. Bouton, J. White.**  
546 2009.  $\alpha 5\beta 1$ -integrin controls ebolavirus entry by regulating endosomal cathepsins. *PNAS*  
547 **106**:8003-8.
- 548 37. **Simmons G., J. D. Reeves, C. C. Grogan, L. H. Vandenberghe, F. Baribaud, J. C.**  
549 **Whitbeck, E. Burke, M. J. Buchmeier, E. J. Soilleux, J. L. Riley, R. W. Doms, P. Bates, and**  
550 **S. Pöhlmann.** 2003. DC-SIGN and DC-SIGNR bind ebola glycoproteins and enhance infection  
551 of macrophages and endothelial cells. *Virology* **305**:115-23.
- 552 38. **Stevens J., O. Blixt, T. M. Tumpey, J. K. Taubenberger, J. C. Paulson, I. A. Wilson.**  
553 2006. Structure and receptor specificity of hemagglutinin from an H5N1 influenza virus. *Science*  
554 **312**:404-10.

- 555 39. **Towner J. S., T. K. Sealy, M. L. Khristova, C. G. Albariño, S. Conlan, S. a Reeder, P.-L.**  
556 **Quan, W. I. Lipkin, R. Downing, J. W. Tappero, S. Okware, J. Lutwama, B.**  
557 **Bakamutumaho, J. Kayiwa, J. a Comer, P. E. Rollin, T. G. Ksiazek, and S. T. Nichol.** 2008.  
558 Newly discovered ebola virus associated with hemorrhagic fever outbreak in Uganda. *PLoS*  
559 *Pathogens* **4**:e1000212.
- 560 40. **Turk B., D. Turk, and V. Turk.** 2000. Lysosomal cysteine proteases: more than scavengers.  
561 *Biochimica et biophysica acta* **1477**:98-111.
- 562 41. **Turk D., M. Podobnik, R. Kuhelj, M. Dolinar, and V. Turk.** 1996. Crystal structures of  
563 human procathepsin B at 3.2 and 3.3 Angstroms resolution reveal an interaction motif between a  
564 papain-like cysteine protease and its propeptide. *FEBS Letters* **384**:211-4.
- 565 42. **Turk D., and G. Guncar.** 2003. Lysosomal cysteine proteases (cathepsins): promising drug  
566 targets. *Acta crystallographica. Section D, Biological Crystallography* **59**:203-13.
- 567 43. **Turk V., B. Turk, and D. Turk.** 2001. Lysosomal cysteine proteases: facts and  
568 opportunities. *The EMBO Journal* **20**:4629-33.
- 569 44. **Vance J. E., and K. B. Peake.** 2011. Function of the Niemann-Pick type C proteins and their  
570 bypass by cyclodextrin. *Current Opinion in Lipidology* 1-6.
- 571 45. **Weidmann, M., E. Mühlberger, F. Hufert.** 2004. Rapid detection protocol for filoviruses.  
572 *Journal of Clinical Virology.* **30**: 94-99.

573 46. **Wong A. C., R. G. Sandesara, N. Mulherkar, S. P. Whelan, and K. Chandran.** 2010. A  
574 forward genetic strategy reveals destabilizing mutations in the Ebolavirus glycoprotein that alter  
575 its protease dependence during cell entry. *Journal of Virology* **84**:163-75.

576

577 **FIGURE LEGENDS**

578 **FIGURE 1: Endosomal cysteine proteases are host factors for filoviruses.**

579 **(A)** Vero cells were incubated in the presence of E-64 (cysteine protease inhibitor, 300  $\mu$ M) or  
580 vehicle (1% DMSO) for 4 hours prior to being challenged by VSV particles encoding GFP  
581 (VSV<sub>GFP</sub>) and bearing GPs from EBOV-May (Z), CIEBOV (CI), SUDV (S), RESTV (R), MARV  
582 (M) or VSV (G). After 24 hours, infectivity [infectious units (i.u.)/ml] was determined by  
583 manually counting GFP positive cells using fluorescence microscopy. Data is mean  $\pm$  s.d. (n=3).  
584 Shown is representative of four independent experiments.

585 **(B)** Vero cells were incubated in the presence of E-64d (cysteine protease inhibitor, 300  $\mu$ M) or  
586 vehicle for 4 hours prior to being challenged by the filoviruses EBOV-May or SUDV (MOI=0.1).  
587 After 3 days, titer was determined by qRT-PCR. Data is mean of three wells  $\pm$  s.d. (n=3).

588 **(C)** Vero cells were incubated in the presence of E-64d (300  $\mu$ M) or vehicle for 4 hours prior to  
589 being challenged by the filoviruses EBOV or MARV (MOI=0.2). After 72 hours, copies of the  
590 viral L gene per cell were determined by qRT-PCR. Data is mean of three wells  $\pm$  s.d. (n=3).

591

592 **FIGURE 2: Function of protease cleaved ebolavirus GPs**

593 **(A)** VSV particles bearing GPs from EBOV-May (Z), CIEBOV (CI), SUDV (S) and RESTV (R)  
594 were incubated with chymotrypsin or reaction buffer alone for one hour. Virus particles were  
595 deglycosylated with PNGase F and analyzed by immunoblot using rabbit anti-GP1 antibody.

596 **(B)** Untreated CHO cells expressing NPC1 or treated with E-64d (300  $\mu$ M) and CHO NPC<sup>-/-</sup> cells  
 597 were challenged by VSV particles encoding luciferase (VSV<sub>luc</sub>) and bearing GPs from uncleaved  
 598 or chymotrypsin cleaved  $\Delta$ Muc EBOV-May (Z), CIEBOV (CI), SUDV (S), or RESTV (R), or  
 599 VSV G. After 16 hours, infection was measured in relative luminescence units (RLU). Data is  
 600 mean  $\pm$  s.d. (n=3).

601 **(C)** SUDV <sub>$\Delta$ TM</sub> GP was cleaved with chymotrypsin (Chy) and binding to late endosome/lysosome  
 602 (LE/LY) membranes from CHO cells expressing NPC1 (right panel) or lacking NPC1 (middle  
 603 panel) determined as described in methods. Bound proteins was analyzed by immunoblot for GP1.  
 604 Uncleaved SUDV <sub>$\Delta$ TM</sub> GP and thermolysin (Thl) cleaved EBOV-May <sub>$\Delta$ TM</sub> GP were included as  
 605 controls. Input proteins are shown in the left panel.

606

607 **FIGURE 3: Cathepsin B is not an essential host factor for all filoviruses.**

608 **(A, B)** Effect of cathepsin B (Cat B) selective inhibitor CA074 on Cat B and cathepsin L (Cat L)  
 609 activity (A) and on infectivity (B) by VSV<sub>GFP</sub> particles bearing filovirus GPs. Vero cells were  
 610 incubated in the presence of increasing concentrations of CA074 (0-80  $\mu$ M) or vehicle for four  
 611 hours prior to cell lysis or the addition of VSV<sub>GFP</sub> particles bearing GPs from MARV (M) or  $\Delta$ Muc  
 612 viruses from EBOV-May (Z), CIEBOV (CI), SUDV (S), or RESTV (R). (A) Cat B and Cat L  
 613 activity were determined by fluorogenic substrates (n=2). (B) After 24 hours, GFP positive cells  
 614 were manually counted using fluorescence microscopy. Infectivities from each of two replicates  
 615 are shown and are representative of four independent experiments. Data is presented as: (infection  
 616 (i.u./ml) of CA074 treated cells) / (infection of vehicle treated cells) X 100%.

617 (C) Wild type and Cat B-deficient (Cat B<sup>-/-</sup> Cat L<sup>+/+</sup>) MEFs were infected with VSV<sub>GFP</sub> particles  
618 bearing GPs from MARV (M) or ΔMuc viruses from EBOV-May (Z), CIEBOV (CI), SUDV (S),  
619 or RESTV (R) or VSV (G). After 24 hours, infectivity was determined as in (B). Data is presented  
620 as: (infection (i.u./ml) of Cat B-deficient cells) / (infection of wild type MEFs) X 100%. Data is  
621 mean +/- s.d. (n=3). Shown is representative of three independent experiments.

622 (D) Wild type and Cat B-deficient MEFs were transfected with expression plasmids encoding  
623 mouse Cat B, mouse Cat L or sham plasmid. After 24 hours, cells were exposed to VSV<sub>GFP</sub>  
624 particles as above. After 24 hours, the percentage of cells infected was determined using flow  
625 cytometry. Data is presented as: (percent infection of transfected Cat B-deficient MEFs) / (percent  
626 infection of wild type MEFs) X 100%. Data is mean +/- s.d. (n=3). Shown is representative of  
627 three independent experiments.

628

629 **FIGURE 4: Cathepsin L is a host factor for Reston and Marburg viruses.**

630 (A, B) Effect of the inhibitor FYdmk on Cathepsin L (Cat L) and Cathepsin B (Cat B) activity (A)  
631 and on infectivity (B) of VSV<sub>GFP</sub> particles bearing filovirus GPs. Vero cells were incubated in the  
632 presence of increasing concentrations of FYdmk (0-10 μM) or vehicle for four hours prior to cell  
633 lysis or the addition of VSV<sub>GFP</sub> particles bearing GPs from MARV (M), ΔMuc viruses from  
634 EBOV-May (Z), CIEBOV(CI), SUDV(S), or RESTV (R). (A) Cat B and Cat L activity were  
635 determined by fluorogenic substrates (n=2). (B) After 24 hours, GFP positive cells were manually  
636 counted using fluorescence microscopy. Infectivities from two replicates are shown and are  
637 representative of four independent experiments. Data is presented as: (infection (i.u./ml) of  
638 FYdmk treated cells) / (infection of vehicle treated cells) X 100%.

639 **(C)** Wild type and Cat B/Cat L-deficient (Cat B<sup>-/-</sup> Cat L<sup>-/-</sup>) MEFs were infected with VSV<sub>GFP</sub>  
 640 particles bearing GPs from MARV (M), ΔMuc viruses from EBOV-May (Z), CIEBOV(CI),  
 641 SUDV(S), and RESTV (R), or VSV (G). Infectivity was determined as in Fig. 3C. Data is mean  
 642 +/- s.d. (n=3). Shown is representative of three independent experiments.

643 **(D)** Cat B/Cat L-deficient MEFs were transfected with expression plasmids encoding mouse Cat  
 644 B, mouse Cat L, both Cat B and Cat L or vector alone. After two days, cells were exposed to  
 645 VSV<sub>luc</sub> bearing GPs from MARV (M) or ΔMuc viruses from EBOV-May (Z), CIEBOV(CI),  
 646 SUDV(S), and RESTV(R) or VSV (G). Twenty-four hours later, infectivity was determined by  
 647 measuring relative luminescence units (RLU) in cell lysates. Data is presented as virus encoded  
 648 (RLU in cathepsin transfected cells) / (RLU of the vector transfected cells). Data is mean +/- s.d.  
 649 (n=6). Shown is representative of four independent experiments.

650

651 **FIGURE 5: Comparison of Cat B requirements for EBOV and RESTV.**

652 **(A)** Virus determinants of Cat B dependence. Vero cells were incubated in the presence of CA074  
 653 (80 μM), E-64 (300 μM) or vehicle (1% DMSO) for 3 hours prior to the addition of VSV<sub>GFP</sub>  
 654 particles bearing ΔMuc GPs from EBOV-May, RESTV, EBOV-May GP1/RESTV GP2, or  
 655 RESTV GP1/ EBOV-May GP2. Infectivity was determined as in Fig. 1A. Data is mean +/- s.d.  
 656 (n=3). Shown are representative of three independent experiments.

657 **(B)** Schematic of ebolavirus EBOV-May GP and location of amino acid residues previously shown  
 658 to mediate resistance to CA074. Inverted triangles identify residues that mediate resistance. Boxes  
 659 identify differences in amino acids between EBOV-May and RESTV. Positions of receptor

660 binding domain (RBD),  $\beta$ -13-14 disordered loop, fusion loop (fl), and heptad repeats 1 and 2 (hr1,  
661 hr2) are indicated. Residue numbering relative to EBOV GP.

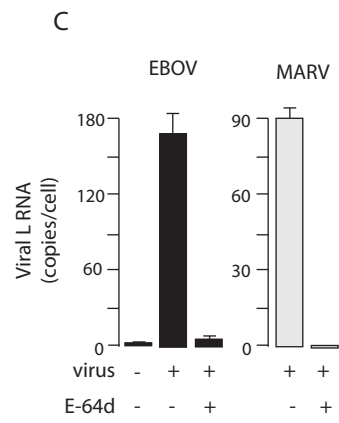
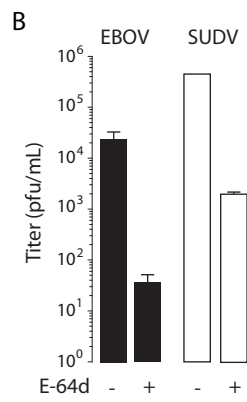
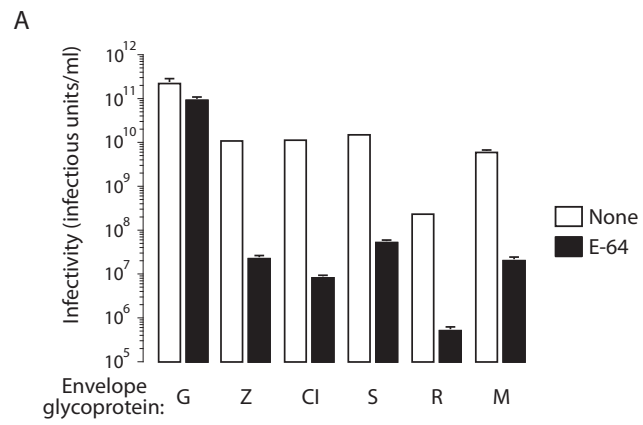
662 **(C and D)** Analysis of D47E and I584L on cathepsin B dependence. The effects of reciprocal  
663 substitutions of residues (D47/E48 and I584/L585) that differ between EBOV-May and RESTV  
664 GP were measured in native and chimeric GPs from A. Infectivity was determined as in Fig. 1A.  
665 Data is mean  $\pm$  s.d. (n=3). Shown are representative of three independent experiments.

666

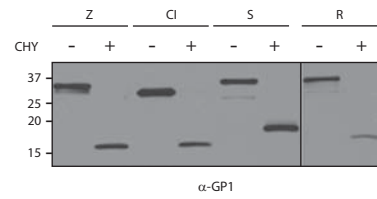
667 **FIGURE 6: Analysis of Cathepsin B dependence of ebolavirus 1995 and Bundibugyo.**

668 **(A)** VSV<sub>GFP</sub> particles pseudotyped with GP from EBOV-May containing E47 or T544 from  
669 EBOV-1995 were prepared and dependence on Cat B analyzed. Vero cells were incubated in the  
670 presence of CA074 (80  $\mu$ M), E-64 (300  $\mu$ M) or vehicle for 3 hours prior to the addition of VSV<sub>GFP</sub>  
671 bearing  $\Delta$ Muc GPs from EBOV-May D47E, I544T or wild type EBOV-May. Infection was  
672 measured as Fig. 1A. Data is mean  $\pm$  s.d. (n=3) Shown is representative of three independent  
673 experiments.

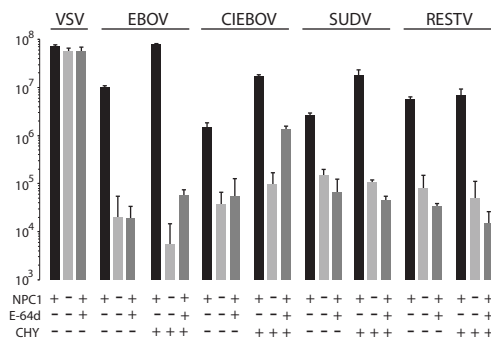
674 **(B)** Effect of Cat B selective inhibitor CA074 on infection by Bundibugyo ebolavirus (BDBV) GP.  
675 Vero cells were incubated in the presence of increasing concentrations of CA074 (0-80  $\mu$ M) or  
676 vehicle for 3 hours prior to the addition of VSV<sub>luc</sub> particles bearing  $\Delta$ Muc GPs from EBOV-May  
677 (Z), BDBV (B) and RESTV (R). After 6 hours, cells were lysed and relative luminescence (RLU)  
678 was measured. Data is presented as virus encoded luciferase activity of (cells treated with CA074)  
679 / (cells treated with vehicle) X 100%. Shown and are representative of three independent  
680 experiments.



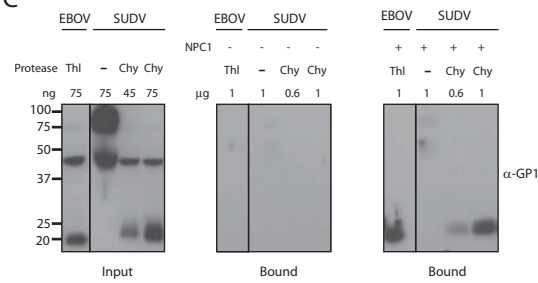
A

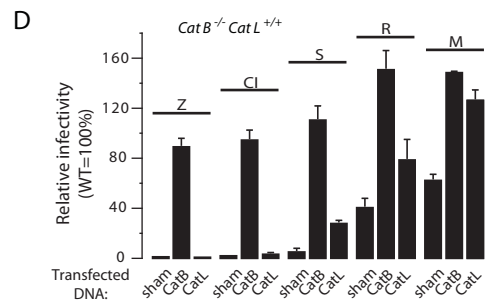
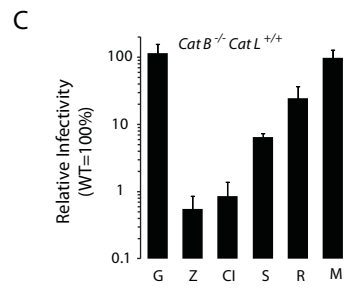
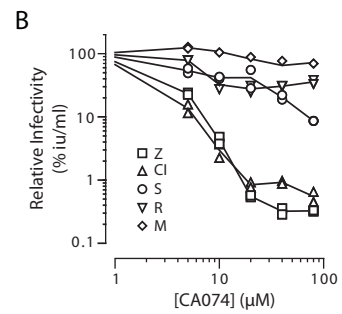
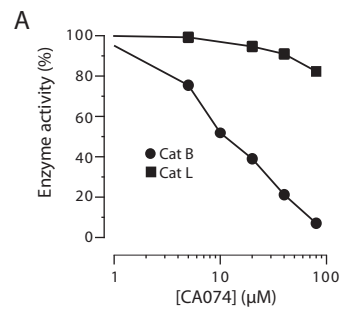


B

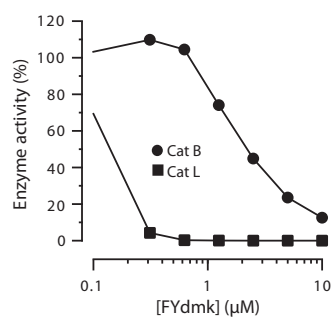


C

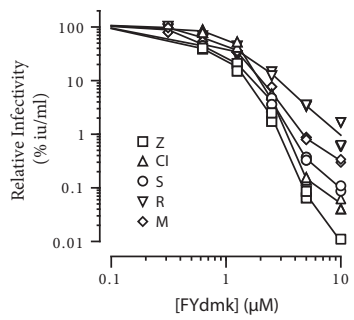




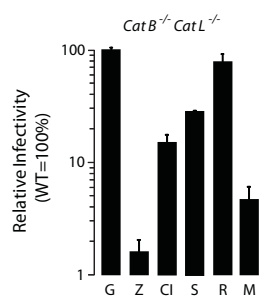
A



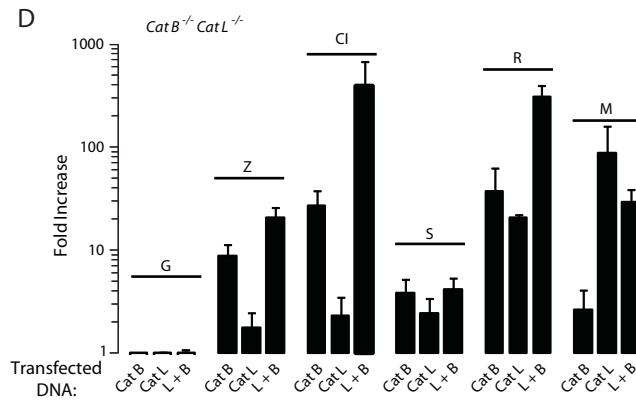
B

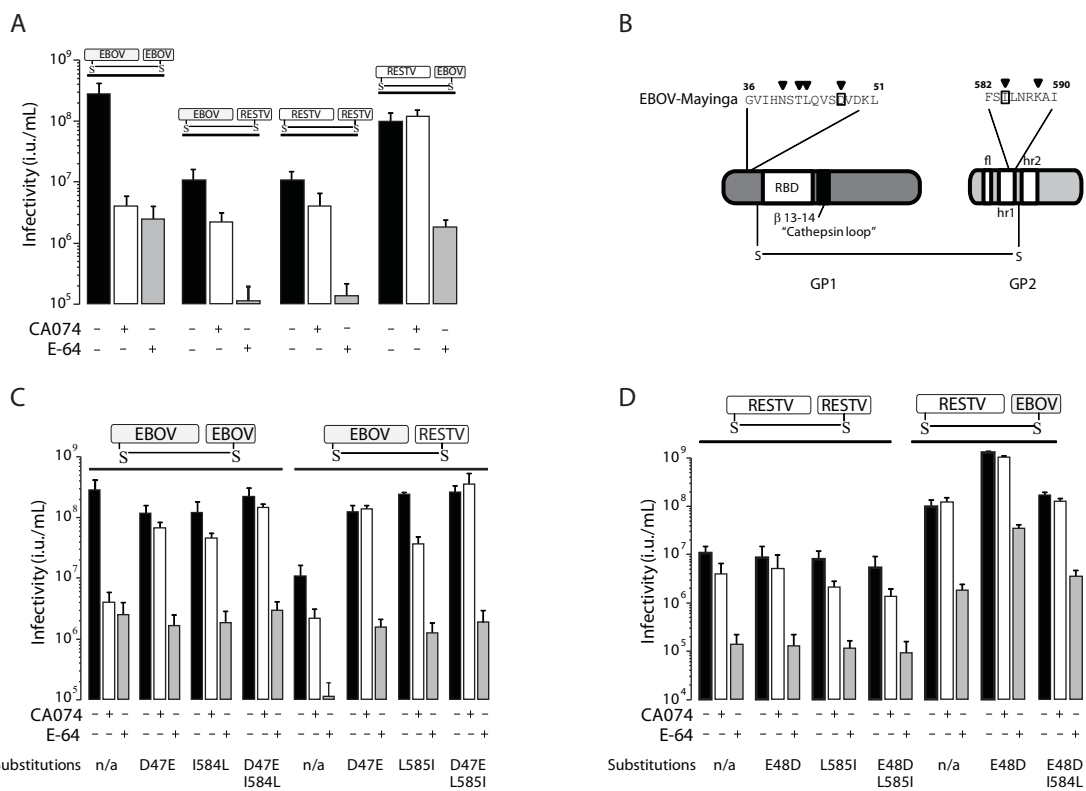


C

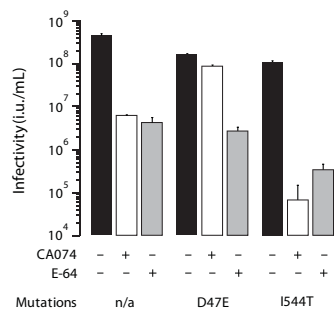


D





A



B

