

Current Biology, Volume 27

Supplemental Information

Structure-Function Studies Link Class II

Viral Fusogens with the Ancestral

Gamete Fusion Protein HAP2

Jennifer Fricke Pinello, Alex L. Lai, Jean K. Millet, Donna Cassidy-Hanley, Jack H. Freed, and Theodore G. Clark

Figure S1, Over-expression of HAP2 in *T. thermophila*. Related to Figure 2.

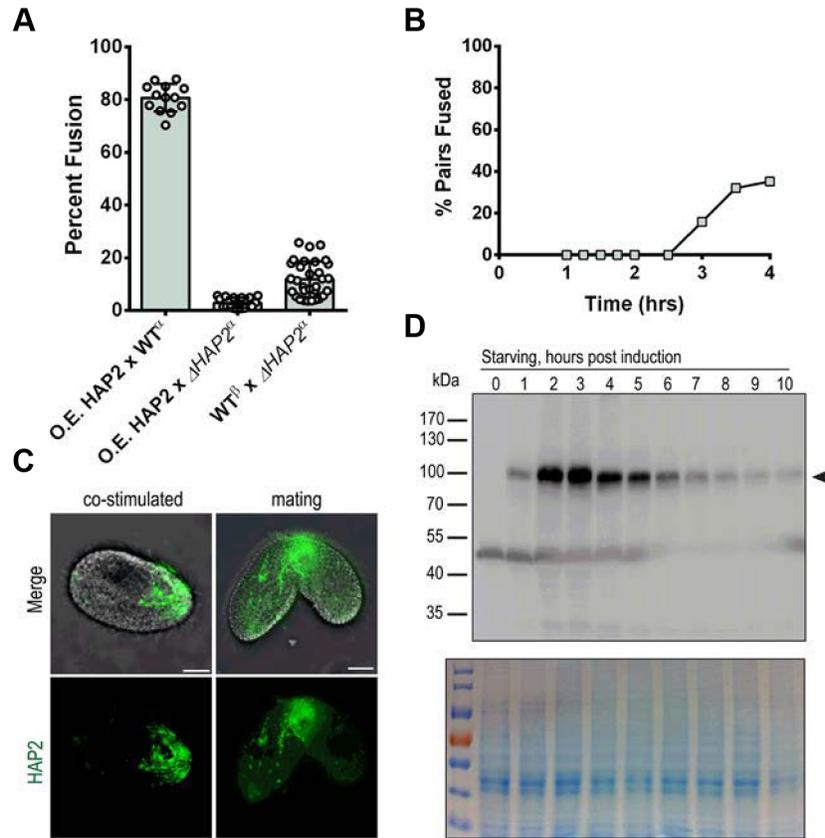


Figure S1 (See also Figure 2). In WT x Δ HAP2 crosses the kinetics of fusion, as well as the final percentage of cells capable of fusion was significantly reduced. To determine whether these effects were due to a reduction in the total amount of HAP2 expressed in a given mating pair, or its pattern of expression in only one mating partner, we over-expressed a full-length HA-tagged version of HAP2 cDNA in a WT partner and mated these cells with a Δ HAP2 knockout strain. We generated the over-expression strain by cloning the tagged HAP2 cDNA into a stable high-copy ribosomal DNA vector and used a robust cadmium-inducible promoter to drive the expression of HAP2 transcripts. Over-expressing cells (designated, O.E. HAP2) were induced with 0.1 μ g/mL CdCl₂ 30 min prior to mating. **(A)** Bar chart showing the percent fusion in crosses between the O.E. HAP2 strain and either a WT or Δ HAP2 partner. The data for WT^β x HAP2^α crosses from Figure 2G is shown for comparison. HAP2 over-expression had no effect on percent fusion in crosses with the WT partner, and slightly diminished fusion in crosses with the Δ HAP2 strain. **(B)** A single experiment showing that the kinetics of fusion in O.E. HAP2 x Δ HAP2 cross. Results from this kinetic experiment were similar to those seen for WT x Δ HAP2 matings (see Figure 1R for comparison). To verify that HAP2 was in fact over-expressed and correctly localized in the O.E. strain we examined these cells by immunofluorescence microscopy and Western blotting following induction with CdCl₂. **(C)** Fluorescence (below) and merged bright field-fluorescence images (above) showing representative single and paired cells from an O.E. HAP2 x WT mating culture fixed 3 h after mixing complementary mating types and immunolabeled with anti-HA antibodies. The co-stimulated cell on the left, and the mating pair on the right showed expression and correct localization of the over-expressed recombinant HAP2 protein at the anterior tip of the co-stimulated cell, and the conjugation junction of mating cells, respectively. **(D)** A time course of recombinant HAP2 expression in unmated O.E. HAP2 cells in 10 mM Tris buffer induced with 0.1 μ g/mL CdCl₂ for 10 hrs as detected by Western blotting. A strong signal just above the expected size of the O.E. HAP2 protein (arrow) was readily detected in cell lysates of the O.E. HAP2 strain beginning at ~ 1 h post induction. It is worth noting that when the same epitope-tagged version of the HAP2 cDNA was placed under the control of the endogenous promoter at the HAP2 locus, mating cultures showed signals by immunofluorescence microscopy (Figure S3) but not by Western blotting suggesting the protein may be highly sensitive to degradation following cellular lysis.

Figure S2, Template-based homology modeling of HAP2. Related to Figure 3.

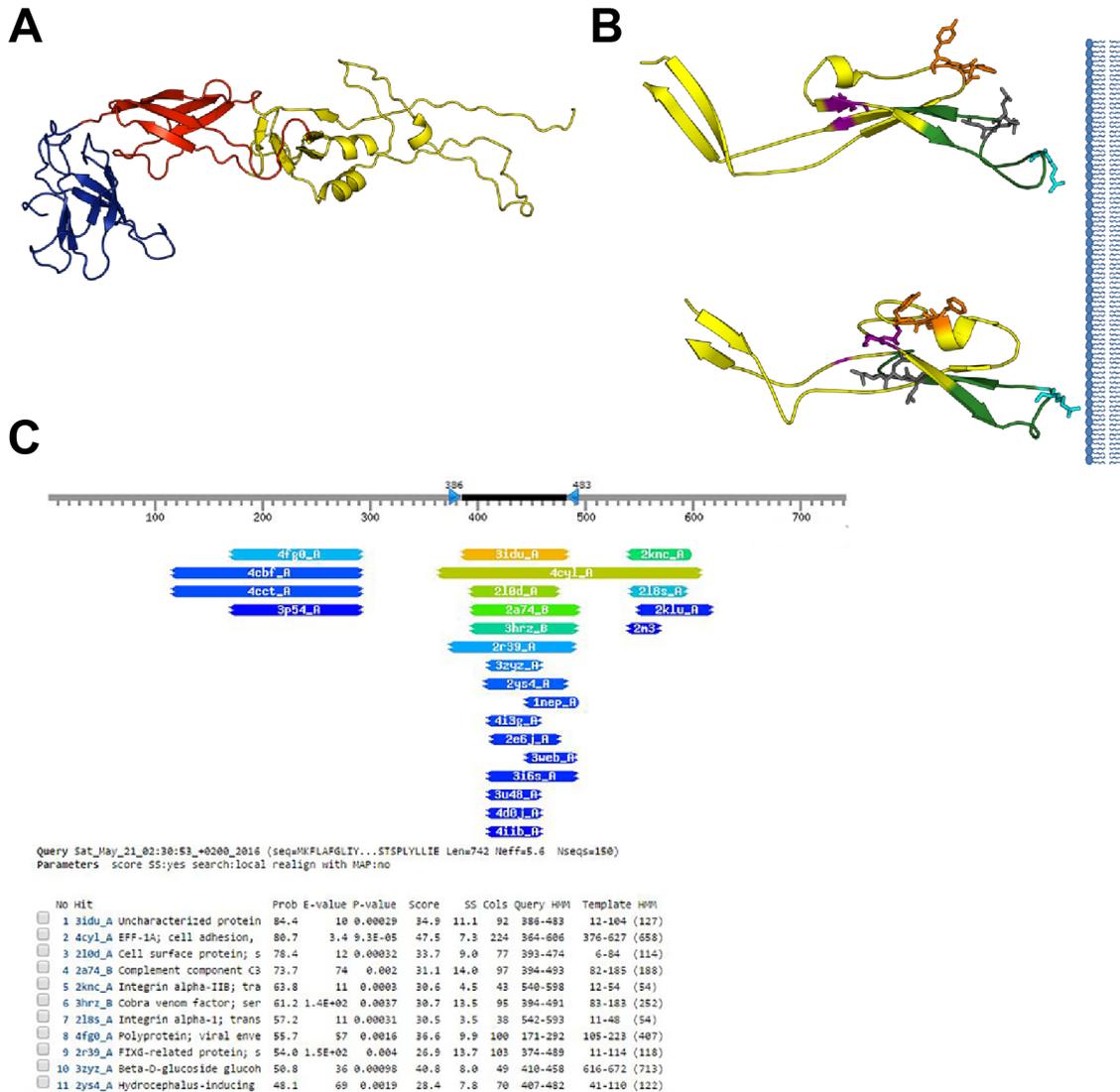


Figure S2 (See also Figure 3). (A) A CPHmodels3.0 generated partial structure of the *T. thermophila* HAP2 ectodomain based on the DENV template PDB ID: 1UZG (aligned residues are 172-478). Domain coloring above follows the convention for class II viral fusogens: domain I is red, domain II is yellow, and domain III is blue. (B) Magnified views of the locations of the tested site-directed mutations within and near the HAP2 fusion loop shown on the partial predicted structures from Phyre2 (top, residues 107-193) and RaptorX (bottom, residues 96-193) and juxtaposed to a cartoon membrane (not drawn to scale). The region of the fusion loop that was truncated is shown in green. Site-directed mutations are shown as sticks with CC147-8 in magenta, FQY131-3 in orange, R164 in cyan, and LNL171-3 in grey. The cysteines predicted to form disulfide bonds with residues CC147-8 are shaded magenta (but not shown as sticks). Depending on loop orientation, which is difficult for template-based modelers to predict, the Phyre2 and RaptorX models show a second loop (containing the FQY131-3 mutation) that might also approach the lipid bilayer. Nevertheless, alteration of the FQY motif in this loop had no effect on the fusogenic activity. (C) A screenshot of HHpred top hit results to the *T. thermophila* HAP2 protein sequence. This was the only template-based modeling tool we tested that identified a homology between HAP2 and EFF-1 (the developmental cell-cell fusogen from *C. elegans*).

Figure S3, Immunofluorescence localization of mutated versions of HAP2. Related to Figure 4.

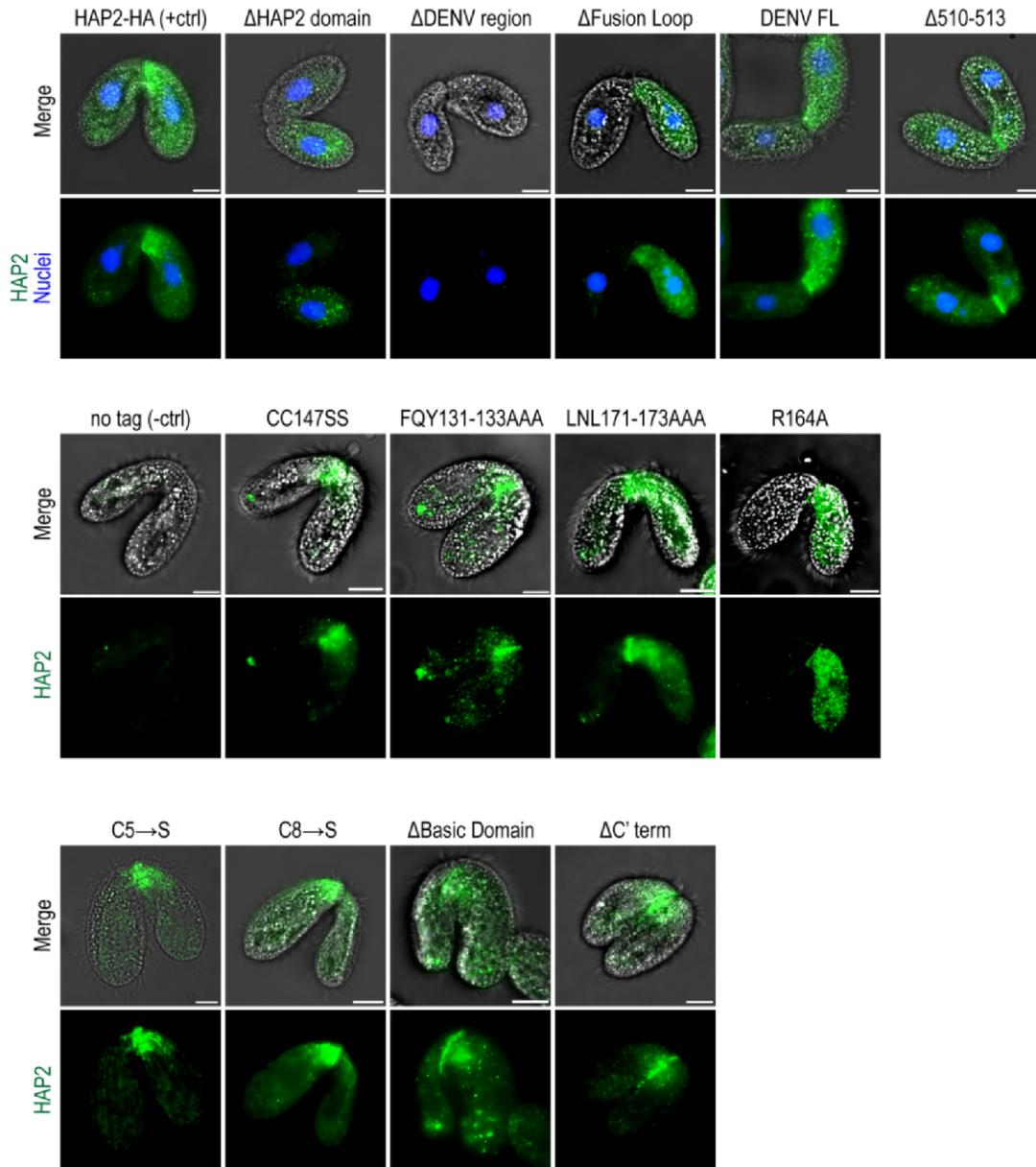


Figure S3 (See also Figure 4). HAP2 localizes to the conjugation junction of *T. thermophila*, a region where cells of complementary mating types adhere and form membrane pores. In each panel, localization of a C'-terminal HA- or FLAG-tagged version of HAP2 is shown in crosses between WT cells and strains harboring mutations/truncations to the *HAP2* coding sequence. In all cases, mutated/truncated gene constructs were targeted to the *HAP2* locus and expressed under the control of the endogenous promoter. Cells were fixed 2.5-5 h after mixing of complementary mating types, then permeabilized and immunolabeled with anti-HA or anti-FLAG antibodies. In some cases, nuclei were labeled 30 min prior to fixation with Hoechst 33258. Each row of paired images shows a merged fluorescence-bright field image (above), and the fluorescence image alone (below), for a representative mating pair from crosses containing the indicated mutated/truncated construct (as labeled above the paired images). No signal was seen in matings between cells that lacked epitope-tagged HAP2. Scale bars are 10 μ m.

Figure S4, HAP2 fusion assays in heterologous systems. Related to Figure 5.

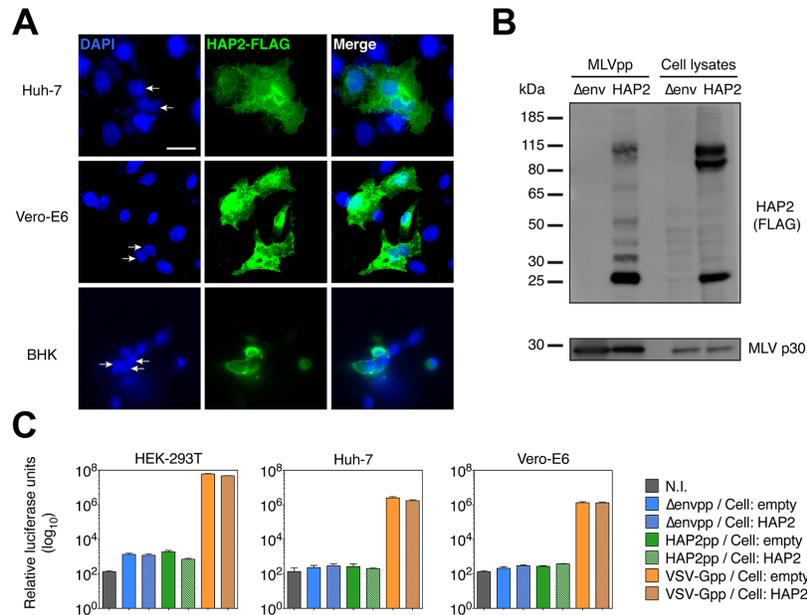


Figure S4 (See also Figure 5). To assess the capacity of *Tetrahymena thermophila* HAP2 to be expressed and mediate membrane fusion events in a heterologous system we carried out immunofluorescence microscopy experiments based on the expression of HAP2 in mammalian cell lines. **(A)** Transfection of human liver (Huh-7), African green monkey (Vero-E6), and baby hamster kidney (BHK) cells with a codon-optimized *T. thermophila* HAP2 gene fused to a 3x FLAG epitope tag resulted in detectable protein expression in all lines tested via immunofluorescence. In this case, transfected cells were fixed, permeabilized and immunolabeled with anti-FLAG antibodies, followed by DAPI staining to label cell nuclei. The HAP2 protein appeared to be expressed on the plasma membrane, as the outline of cells, including the filopodia and membrane ruffles of the Huh-7 cells, could be well delineated after labeling. Although large syncytia were not seen, numerous examples of multi-nucleated HAP2-expressing cells (arrows, left-hand panels) were visible from transfection of all three cell lines. Scale bar is 25 μ m. **(B)** Pseudotyped particles (pp) are enveloped virus particles (often derived from retroviruses) that harbor the fusion proteins of a heterologous virus on their surface. These particles are excellent surrogates of native viruses to study virus entry and fusion [S1]. To determine whether HAP2 function was sufficient to fuse viral envelopes to cells we generated murine leukemia virus (MLV) retroviral pseudotyped particles that incorporated heterologous HAP2 (HAP2pp). Panel (B) shows western blotting analysis of HAP2 protein expression in transfected cells and HAP2 incorporation into pseudotyped particles derived from those cells. HEK-293T cells were co-transfected to produce either *T. thermophila* HAP2-3xFLAG pseudotyped particles (HAP2pp) or no envelope protein control particles (Δ envpp). After supernatant harvesting, the pseudotyped particle producer cells were lysed and analyzed by Western blot along with the concentrated pseudotyped particles. HAP2-3xFLAG was detected using an anti-FLAG tag antibody, and MLV p30 capsid protein (loading control) was detected using the anti-MLV p30 capsid antibody (4B2). In both HAP2-transfected cell lysates and in HAP2pp, the HAP2 protein could be detected. In cell lysates, bands were detected at \sim 115, \sim 100, and \sim 25 kDa, whereas in HAP2pp only the \sim 115 and 25 kDa bands, were seen, indicating preferential incorporation of certain forms of HAP2 into the particles. Several bands of intermediate sizes were also observed suggesting HAP2 sensitivity to proteolysis. **(C)** The pseudotyped particles in (B) contained a luciferase gene that integrates into the target cell's genome if successful viral membrane fusion and entry has occurred. Assaying for luciferase activity in target cells after application of particles allowed for determination of the infectivity, and thus fusogenicity, of each type of pseudotyped particle analyzed. Cell lines used to test the infectivity of these particles were HEK-293T, Huh-7, and Vero-E6 cells that were either transfected with an empty vector, or a vector designed to express HAP2-3xFLAG (as in panel A). Twenty-four hours post transfection, HAP2pp, Δ envpp, and VSV-G pseudotyped particles (positive control particles pseudotyped with the VSV-G envelope glycoprotein) were used to inoculate cells. Seventy-two hours post infection, cells were lysed and luciferase activity was determined using a luminometer. Note the relatively low levels of luciferase activity in the control Δ envpp and HAP2pp infected samples (\sim 10^2 - 10^3 relative luciferase units, RLU).

Table S1, Phyre2 batch processing results. Related to Figure 3.

Organism, GenBank accession #^a	CII viral hits^b	Non viral envelope hits
<i>Gonium pectorale</i> , BAO57178.1	11 (61.3%, 2OF6, WNV) 14 (59.4%, 3UAJ, DENV)	1 (92.0%, 4FXK, complement c4-a) 2 (82.4%, 3BGA, beta-galactosidase) 3 (82.3%, 1CZD, DNA POL protein g45)
<i>Chlorella variabilis</i> , XP_005851393.1	1 (95.5%, 4B03, DENV) 2 (94.9%, 1UZG, DENV) 4 (94.7%, 1SVB, TBEV)	9 (89.0%, 2X41, beta-glucosidase) 10 (85.7%, 2E6J, papd-like domain, hydin) 11 (84.5%, 3QBT, INPP5 ocr1-1)
<i>Volvox carteri</i> , XP_002952884.1	5 (42.9%, 2OF6, WNV) 11 (23.6%, 4CBF, DENV)	1 (58.0%, 2R39, protein from <i>V. parhaemolyticus</i>) 2 (51.4%, 2E6J, papd-like domain, hydin) 3 (46.8%, 2RNR, TFiih complex p62)
<i>Chlamydomonas reinhardtii</i> , ABO29824.2		1 (88.5%, 4FXK, complement c4-a) 2 (83.9%, 2E6J, papd-like domain, human hydin) 3 (81.4%, 1JZ8, beta-galactosidase)
<i>Helicosporidium sp.</i> , KDD77085.1	3 (36.9%, 2OF6, WNV) 4 (36.9%, 4CBF, DENV)	1 (39.7%, 2M0G, splicing factor 1) 5 (23.9%, 1M2s, K-channel blocking toxin bmtx3) 6 (23.0%, 4QE0, duf5043 from <i>B. uniformis</i>)
<i>Coccomyxa subellipsoidea</i> , XP_005651045.1	1 (90.9%, 3UAJ, DENV) 3 (80.9%, 2OF6, WNV) 6 (75.1%, 1UZG, DENV)	4 (79.0%, 2E6J, papd-like domain, hydin) 5 (76.8%, 2X41, beta-glucosidase) 13 (66.0%, 4FXK, complement c4-a)
<i>Lilium longiflorum</i> , BAE71142.1	1 (96.3%, 3UAJ, DENV) 2 (94.3%, 1OK8, DENV) 3 (93.3%, 1SVB, TBEV)	9 (84.7%, 4IID, beta-glucosidase 1) 11 (77.9%, 2E6J, papd-like domain, hydin) 20 (64.9%, 2L0D, <i>m. acetivorans</i> protein)
<i>Zea mays</i> , NP_001307741	8 (59.0%, 3UAJ, DENV)	1 (91.3%, 2XC8, <i>B. subtilis</i> spp1 phage) 2 (80.7%, 2L0D, <i>m. acetivorans</i> protein) 3 (79.6%, 4FXK, complement c4-a)
<i>Arabidopsis thaliana</i> , NP_192909.2	1 (92.9%, 1URZ, TBEV) 2 (91.5%, 1OK8, DENV) 6 (89.6%, 1SVBA, TBEV)	3 (90.0%, 2KL6, cardb domain of <i>p. furiosus</i>) 4 (89.8%, 1JZ8, beta-galactosidase) 8 (84.6%, 4FXK, complement c4-a)
<i>Cyanidioschyzon merolae</i> , XP_005536505.1		1 (85.6%, 2X41, beta-glucosidase) 6 (69.0%, 3QIS, INPP5 ocr1-1) 9 (65.8%, 2E6J, pap-d like domain, hydin)
<i>Galdieria sulphuraria</i> , XP_005708101.1		1 (92.7%, 1W8O, bacterial sialidase) 2 (89.6%, 4FXK, complement c4-a) 3 (87.4%, 3ZZ1, beta-d-glucoside glucohydrolase)
<i>Trypanosoma brucei</i> , XP823296.1	1 (96.7%, 1UZG, DENV) 2 (96.3%, 1SVB, TBEV) 5 (95.1%, 2OF6, WNV)	11 (86.6%, 2X41, beta-glucosidase) 14 (60.6%, 2MI2, protein translocase protein tatb) 17 (56.7%, 2E6J, papd-like domain, hydin)
<i>Trypanosoma cruzi</i> , XP_814894.1	1 (95.1%, 3UAJ, DENV) 2 (92.7%, 2OF6, WNV) 3 (91.6%, 1SVB, TBEV)	9 (81.7%, 1JZ8, beta-galactosidase) 15 (57.9%, 3QBT, INPP5 ocr1-1) 19 (45.7%, 2XC8, <i>B. subtilis</i> spp1 phage)
<i>Strigomonas culicis</i> , EPY22600.1	1 (91.1%, 1UZG, DENV) 2 (90.4%, 1SVB, TBEV) 3 (90.4%, 2OF6, WNV)	5 (86.4%, 1JZ8, beta-galactosidase) 6 (80.7%, 2E6J, papd-like domain, hydin) 12 (60.9%, 2XC8, <i>B. subtilis</i> spp1 phage)

<i>Angomonas deanei</i> , EPY38446.1	3 (83.9%, 2OF6, WNV)	1 (88.1%, 1JZ8, beta-galactosidase) 2 (85.9%, 2XC8, <i>B. subtilis</i> spp1 phage)
<i>Phytomonas</i> sp. Isolate EM1, CCW64758.1	1 (94.9%, 1SVB, TBEV) 2 (93.8%, 2OF6, WNV) 3 (92.0%, 1UZG, DENV)	8 (86.2%, 1JZ8, beta-galactosidase) 12 (73.1%, 4FCK, complement c4-a) 16 (62.3%, 2E6J, papd-like domain, human hydin)
<i>Leishmania major</i> , XP_003722443.1	1 (93.3%, 1UZG, DENV) 4 (88.7%, 1SVB, TBEV) 5 (86.3%, 2OF6, WNV)	7 (81.6%, 2E6J, papd-like domain, human hydin) 8 (76.5%, 2X41, beta-glucosidase) 11 (71.5%, 3QBT, INPP5 ocr1-1)
<i>Naegleria gruberi</i> , XP_002674350.1	1 (93.3%, 3UAJ, DENV) 4 (90.3%, 1SVB, TBEV) 6 (88.6%, 2OF6, WNV)	9 (73.1%, 2KNC, integrin alpha-iib) 12 (66.8%, 2KL6, cardb domain of p. furiosus) 13 (61.7%, 2V5Y, tyrosine-protein phosphatase mu)
<i>Physarum polycephalum</i> , BAE71144.1		1 (86.3%, 1JZ8, beta-galactosidase) 2 (85.4%, 4FXK, Complement c4-a alpha chain) 3 (73.3%, 2V5Y, tyrosine-protein phosphatase mu)
<i>Dictyostelium fasciculatum</i> , XP_004359139.1	19 (39.6%, 4B03, DENV)	1 (88.2%, 1JZ8, beta-galactosidase) 8 (59.1%, 2JE8, beta-mannosidase) 9 (58.4%, 2XC8, <i>B. subtilis</i> spp1 phage)
<i>Acanthamoeba castellanii</i> , XP_004341525.1	1 (96.1%, 3UAJ, DENV) 2 (95.3%, 2OF6, WNV) 3 (94.8%, 1SVB, TBEV)	11 (72.8%, 1JZ8, beta-galactosidase) 13 (67.6%, 2E6J, papd-like domain, human hydin) 16 (64.8%, 2RNR, TFiih complex p62)
<i>Theileria parva</i> , XP_764209.1		1 (60.7%, 1JZ8, beta-galactosidase) 6 (35.1%, 3QBT, INPP5 ocr1-1) 8 (29.1%, 4FXK, complement c4-a)
<i>Toxoplasma gondii</i> , EPT31063.1		1 (81.2%, 1JZ8, beta-galactosidase) 2 (73.2%, 2MI2, translocase protein tatb) 3 (55.1%, 2VRS, capsid protein avian reovirus)
<i>Plasmodium berghei</i> , XP_676900.1		1 (57.9%, 1JZ8, beta-galactosidase) 3 (32.5%, 1A87, colicin n) 8 (16.4%, 2XC8, <i>B. subtilis</i> spp1 phage)
<i>Plasmodium falciparum</i> , XP_001347424.1	2 (37.7%, 3J2W, CHIKV)	1 (56.2%, 1JZ8, beta-galactosidase) 4 (37.2%, 4FXK, complement c4-a) 5 (34.7%, 1A87, colicin n)
<i>Oxytricha trifallax</i> , EJY77656.1		1(98.7%, 4CGK, protein PCSB from <i>S.pneumoniae</i>) 2(98.4%, 3VKG, dynein heavy chain) 5(97.8%, 4L1B, PI3K regulatory subunit alpha)
<i>Tetrahymena thermophila</i> , KJ629172	1 (94.9%, 1UZG, DENV) 3 (89.7%, 2OF6, WNV) 7 (78.4%, 1SVB, TBEV)	8 (76.5%, 1JZ8, beta-galactosidase) 10 (55.6% 4FXK, complement c4-a) 11 (52.5%, 2E6J, papd-like domain, human hydin)
<i>Ichthyophthirius multifiliis</i> ^c	1 (95.2%, 3UAJ, DENV) 3 (91.1%, 1SVB, TBEV) 9 (78.7%, 2OF6, WNV)	4 (84.7%, 1JZ8, beta-galactosidase) 14 (48.2%, 2KL6, cardb domain of p. furiosus) 15 (47.4%, 2E6J, papd-like domain, human hydin)
<i>Paramecium tetraurelia</i> , XP_001431224.1	1 (87.3%, 3UAJ, DENV) 3 (76.9%, 2OF6, WNV) 4 (72.9%, 1UZG, DENV)	2 (77.3%, 2MKV, Na/K-transporting atpase) 5 (72.2%, 2E6J, papd-like domain, human hydin) 7 (65.4%, 2JO1, phospholemman)

<i>Capsaspora owczarzaki</i> , XP_004343268.1		1 (84.7%, 2XC8, <i>B. subtilis</i> spp1 phage) 2 (82.3%, 1JZ8, beta-galactosidase) 3 (82.2%, 4FXK, complement c4-a)
<i>Salpingoeca rosetta</i> , XP_004989263.1	1 (79.5%, 3UAJ, DENV) 5 (64.6%, 2OF6, WNV) 11 (51.9%, 1URZ, TBEV)	3 (68.1%, 2PBD, phosphoprotein) 4 (66.9%, 2LFT, human prion protein with e219k) 7 (60.1%, 1LNZ, spo0b-associated gtp-binding protein)
<i>Monosiga brevicollis</i> , XP_001746497.1	1 (96.2%, 1UZG, DENV) 2 (95.0%, 4B03, DENV) 3 (94.7%, 3UAJ, DENV)	11 (63.2%, 1QK6, huwentoxin-i) 15 (33.8%, 2IEC, protein in <i>m. kandleri</i>) 16 (33.0%, 2V5Y, tyrosine-protein phosphatase mu)
<i>Nematostella vectensis</i> , XP_001628495.1	1 (96.3%, 1UZG, DENV) 2 (95.5%, 1SVB, TBEV) 4 (95.0%, 2OF6, WNV)	11 (78.6%, 2X41, beta-glucosidase) 15 (52.6%, 2E6J, papd-like domain, human hydin) 16 (42.0%, 2R39, protein from <i>V. parahaemolyticus</i>)
<i>Hydra vulgaris</i> , ABN45755.1	3 (73.0%, 2OF6, WNV) 6 (69.2%, 1UZG, DENV)	1 (90.3%, 1JZ8, beta-galactosidase) 5 (70.7%, 2E6J, papd-like domain, human hydin) 13 (49.9%, 4AK2, heparin-binding protein)
<i>Tribolium castaneum</i> , EFA06462.1	1 (62.7%, 1SVB, TBEV) 10 (30.2%, 1UZG, DENV)	2 (48.7%, 3MU3, chicken md-1) 3 (43.4%, 4CCV, histidine-rich glycoprotein) 4 (41.7%, 2CG7, fibronectin)
<i>Acyrtosiphon pisum</i> , XP_003245993.2	3 (36.6%, 1UZG, DENV) 6 (32.7%, 1SVB, TBEV)	1 (52.4%, 2YS4, papd-like domain, human hydin) 2 (47.1%, 1XSZ, GEF ralf) 9 (27.6%, 3U6X, phage tp901-1 baseplate tripod)
<i>Capitella teleta</i> , ELU07639.1	1 (77.6%, 1SVB, TBEV) 2 (75.0%, 2OF6, WNV) 3 (74.4%, 1UZG, DENV)	7 (63.0%, 4GWM, hydrolase, promeprin beta) 8 (53.0%, 3MU3, chicken md-1) 9 (47.0%, 3DUE, periplasmic protein from <i>B. vulgatus</i>)
<i>Pediculus humanus corporis</i> , XP_002429972.1		1 (87.1%, 1GM6, boar salivary lipocalin) 2 (84.0%, 1EW3, allergen equ c1) 3 (82.5%, 3MU3, chicken md-1)
<i>Drosophila melanogaster</i> , NP_001034068.2		1 (55.8%, 1JRJ, exendin-4) 2 (45.1%, 1D0R, glucagon-like peptide) 3 (33.8%, 2CCT, zinc finger domain of DnaJ)
<i>Saccoglossus kowalevskii</i> , XP_006821859.1		1 (50.4%, 1GYG, alpha-toxin from <i>C. perfringens</i>) 2 (48.0%, 3DKB, tumor necrosis factor a20) 4 (24.7%, 1TM9, protein from <i>Mycoplasma genitalium</i>)
<i>Apis mellifera</i> , XP_006565646.1		1 (81.1%, 3QBT, INPP5 ocr1-1) 4 (71.6%, 3MU3, chicken md-1) 11 (50.3%, 2F61, beta-glucosidase)

^aThe name of HAP2 containing organisms (top) and GenBank accession numbers of their respective HAP2 orthologs (bottom). The identities of HAP2 orthologs were based on the findings of previous studies [S2].

^bThe Phyre2 hits to each HAP2 ortholog are listed as Hit Rank (Confidence%; PDB ID; name of protein). The top three ranking hits (out of 20 total) to different templates in each category (viral/non-viral) are shown. A hit rank of 1 is the best-scoring template from which the predicted structure was built. Viral envelope protein hits are abbreviated based on the name of the virus from which they were derived (DENV = Dengue Virus E glycoprotein; TBEV = Tick Borne Encephalitis Virus envelope glycoprotein; WNV = West Nile Virus Envelope glycoprotein; and, CHIKV = Chikungunya virus envelope protein). All species that had a hit to a viral envelope protein structure are shaded grey.

^cThe hits to the *Ichthyophthirius multifiliis* HAP2 ortholog were determined through the Phyre2 processing portal.

Supplemental Experimental Procedures

***Tetrahymena* strains and culture conditions.** *Tetrahymena thermophila* strains were obtained from the Tetrahymena Stock Center, Cornell University (<https://tetrahymena.vet.cornell.edu/>). All established and newly created cell lines are described in the table labeled *Tetrahymena thermophila* strains. For routine growth, cells were incubated at 30°C in NEFF medium (0.25% proteose peptone, 0.25% yeast extract, 0.5% glucose, 33.3 μM FeCl₃) on a platform shaker at ~100 rpm. For mating studies, log phase cells of different mating types were starved in 10 mM Tris buffer (pH 7.5) for up to 48 h, then mixed in equal numbers to a final concentration of 2×10^5 cells/mL at 30°C. For somatic (macronuclear) transformation, target cells were grown to late log phase ($\sim 1 \times 10^6$ cells/mL) in NEFF medium and starved overnight in 10 mM Tris buffer (pH 7.5) for 24 h at $\sim 2 \times 10^5$ cells/mL prior to biolistic transformation and drug selection (see below, *T. thermophila* strain construction).

***T. thermophila* strain construction.** Cell lines used for the creation of HAP2 mutations/truncations at the endogenous *T. thermophila* HAP2 gene locus are designated Δ HAP2-428 clone 5 and Δ HAP2-427 clone 6 and are derivatives of the heterokaryon strains CU428.2 and CU427.4 [S3]. These strains express mating types (VII and VI respectively) and lack the entire HAP2 coding sequence in the macronucleus [S4]. Relevant gene constructs were introduced into the knockout cell lines via biolistic bombardment (see below) and stable drug-resistant transformants were selected by growth in cycloheximide following homologous recombination at the HAP2 locus.

Mutant HAP2 gene constructs were prepared using either overlap PCR, or in some cases, site-directed mutagenesis with a Q5[®] Site-Directed Mutagenesis kit (New England BioLabs). A full list of PCR primers and their use in gene construction is provided in the table labeled, PCR primers. In all cases, epitope tags (either HA or Flag-His) were added to the 3' ends of the HAP2 cDNA constructs by PCR to permit localization of the recombinant gene products. All PCR reactions were carried out with Phusion High-Fidelity *Taq* DNA Polymerase (ThermoFisher). HAP2 PCR products were gel purified and cloned into a previously constructed pCR[™]4Blunt-TOPO[®] plasmid vector backbone (ThermoFisher), which had been modified to contain ~1000 bp of 5' and 3' flanking sequences from the *T. thermophila* HAP2 gene [S4]. *Bam*HI and *Kpn*I restriction sites situated between the HAP2 5' and 3' flanks were used for directional cloning of coding sequences for the mutated/truncated HAP2 gene products [S4]. The modified vector also contained a pHrp129-B cycloheximide resistance cassette within the 3' flanking sequence 387 bp downstream of the *Kpn*I restriction site. Purified HAP2 PCR products and vector DNA were digested with *Bam*HI and *Kpn*I and ligated with T4 DNA ligase (New England BioLabs,) prior to transformation and amplification in *E. coli* 10G competent cells (Lucigen). After transformation, all HAP2 plasmid DNA sequences were verified by sequencing (Cornell Biotechnology Resource Center) to confirm correct gene construction.

For transformation into *T. thermophila*, plasmid DNA harboring relevant inserts was purified, linearized by digestion with *Eco*RI, and introduced into the macronucleus of Δ HAP2 parental strains via biolistic bombardment [S5] using a PDS-1000/He Biolistic Particle Delivery System (Bio-Rad). Positive transformants were selected in NEFF medium containing 25 μg/mL cycloheximide, and supplemented with 1.25 μg/mL Fungizone, 250 μg/mL Streptomycin, and 250 μg/mL Penicillin G. Transformed clones were then pushed to complete macronuclear replacement for the target construct via growth in NEFF medium containing increasing concentrations of cycloheximide (up to 50 μg/mL). Genomic DNA was extracted with phenol:chloroform:iso-Amyl alcohol (25:24:1, VWR), PCR amplified and the resulting products analyzed by agarose gel electrophoresis and, in some cases, Sanger sequencing at the Cornell Biotechnology Resource Center.

In addition to creating strains in which mutant gene constructs were targeted to the endogenous HAP2 locus, a cell line that over-expresses HAP2 transcripts in an otherwise wild type HAP2 background was generated using the stable, high-copy ribosomal DNA vector pTRAS (Tetragenetics). For this strain, a full-length HAP2 cDNA was amplified with the primer pair Hap2BamHIfor/GSP2KpnIrev (see PCR primers table). The resulting PCR product was then digested with *Bam*HI and *Kpn*I, gel purified, and ligated to a *Kpn*I-digested fragment of synthetic DNA containing the following multi-sequence tag: *Kpn*I restriction site; PreScission S protease cleavage sequence; Streptavidin Binding Peptide sequence; and, 6× His tag. The ligation product was then purified and subjected to PCR for the further addition of a 3' HA tag and *Sac*I restriction site using the primer pairs BamHIHAP2for/HA-tag Rev (see PCR primers table). This PCR product was gel purified and restricted with *Bam*HI and *Sac*I for ligation into the shuttle vector pTIEV4 (Tetragenetics). The HAP2-tags-pTIEV4 plasmid was amplified in *E. coli* 10G (Lucigen), purified, and then cut at a unique *Not* I site. The insert was then sub-cloned into pTRAS, downstream of a cadmium-inducible promoter from the *MTT1* gene of *T. thermophila*[S6]. Finally, the recombinant vector was biolistically transformed into 8-10 h mating cultures of *Tetrahymena* strains CU428 x

B2086 as described above. Macronuclear transformants were selected by growth in NEFF medium containing increasing concentrations of paromomycin (up to 800 $\mu\text{g}/\text{mL}$) and then frozen in liquid nitrogen [S7].

Flow cytometry assays for cell-cell fusion. Cells of different mating types were grown and placed in starvation medium (10 mM Tris buffer, pH 7.5) for 24 h to ensure all cells had completed asexual division and were developmentally arrested [S8]. Cells were then washed once by centrifugation at $400 \times g$ and resuspended in $0.1 \times \text{PBS}$ prior to labeling. All labeling reactions and centrifugation steps were carried out in 15 mL glass conical tubes. One mL of $0.1 \times \text{PBS}$ containing 7×10^6 cells was combined with 1 mL of the same buffer containing either 20 μM CFSE (Affymetrix eBioscience) or 10 μM CTFR (Life Technologies). These dyes freely enter cells and become reactive with free amines on proteins and other molecules following cleavage by cytosolic esterases [S9]. Cell suspensions were incubated in the dark for 5 min at room temperature (RT), or 15 min at 30°C for CFSE and CTFR, respectively. Ten mL of NEFF media was then immediately added to quench excess unbound label, and cells were washed and resuspended in 10 mL of 10 mM Tris (pH 7.5). Cells were maintained overnight in the dark at 30°C , and the following day, washed again (10 mL of 10 mM Tris) and counted. Equal numbers of cells of each mating type were then combined in 100 mm \times 10 mm petri dishes ($0.5\text{--}2 \times 10^6$ total cells/dish), and allowed to mate for 16-20 h at 30°C in a darkened incubator. Following mating, exconjugant cells were centrifuged ($350\text{--}400 \times g$) and fixed with IC Fixation buffer (Affymetrix eBioscience) in the dark for 20 min at RT, then resuspended in $1 \times \text{PBS}$ containing 0.3% BSA prior to acquisition on a BD FACSCantoTM II Flow Cytometer.

For each mating reaction a minimum of 30,000 events were acquired. Data were analyzed using FlowJo software (FlowJo LLC). Unmated, single-labeled control cell cultures were fixed at the same time as mating cultures and subjected to flow cytometry as above. Labeled populations from the unmated cultures served as guides for drawing gates around double-labeled populations from experimental mating cultures. The starved, unmated populations were also used to estimate fluorescence loss due to “co-stimulation” (Figure 2) after it became clear these populations had higher fluorescence intensity than the parallel, single-labeled populations in mating cultures. Bar charts displaying the frequency of “Mid”-fluorescence events (percent of cells that had undergone fusion) in each cross were generated using Prism Software (GraphPad Inc.).

Immunofluorescence microscopy. Wild type cells (*T. thermophila* strain CU428.2 [mating type VII]) and cell lines harboring tagged HAP2 constructs (strain CU427.4, mating type VI) were cultured separately, starved and then combined to initiate mating as described above. At time points indicated in the figure legends, mating pairs were washed in 20 mM HEPES buffer, pH 7.4, then fixed for 20 min at room temperature by gentle addition of IC Fixation Buffer (Affymetrix eBioscience) at a ratio of 1:1 with resuspended cells. In some cases, Hoechst 33258 (Invitrogen) was added to the media 30 min prior to fixation for localization of macro- and micronuclei (Fig. 1A-D; I-L). All centrifugation steps were carried out at $350 \times g$. After fixation, cells were again centrifuged and resuspended in $1 \times \text{Permeabilization Buffer}$ (Affymetrix eBioscience), blocked in PBS containing 3% BSA, and incubated overnight at 4°C with the addition of either mouse anti-HA (anti-HA.11, BioLegend, formerly Covance) or rabbit anti-FLAG (Rockland, Inc.) antibodies at a dilution of 1/1000, followed by a 1 h incubation in secondary Alexa 488-conjugated goat anti-mouse and goat anti-rabbit antibodies respectively, at a dilution of 1/1000 (Life Technologies).

For immunofluorescence analyses of HAP2-3xFLAG protein expression in mammalian cell lines, cells were washed in PBS buffer, and fixed for 1 h in PBS containing 4% paraformaldehyde. Cells were then permeabilized with Triton X-100, blocked with normal goat serum, and immunolabeled using first an anti-FLAG antibody (FLAG M2 antibody, Sigma), and next, a goat anti-mouse Alexa488 antibody (Life Technologies) at dilutions of 1/1000 and 1/500 respectively. Upon mounting, nuclei were stained by addition of Fluoromount G with DAPI (Electron Microscopy Sciences).

In both cases (i.e. mating *T. thermophila*, and HAP2-transfected mammalian cell lines), slides were analyzed for expression and localization of HAP2 using either the $63\times$ or $100\times$ objectives on a Zeiss Axio Imager M1 microscope equipped with an AxioCamMR3 camera.

Measurement of cell-cell pairing and membrane fusion kinetics. Cells of different mating types were labeled and mixed together to initiate mating as described above. Samples from the mating cultures were then collected and fixed at the indicated time points within the first 4 h after mixing. Cells were observed under phase and fluorescence optics to determine the percent of cells in pairs (% pairing), as well as the percentage of pairs that had visibly exchanged fluorescent content (% fusion). For % pairing, at least 100 “subjects” (pairs or single cells) were counted for each time point and the percentage calculated as the number of cells in pairs over the total number of cells counted multiplied by 100. For % fusion, at least 50 pairs were counted for each time point with their fusion

status determined based on visual detection of fluorescent tracer in only one (not fused) versus both (fused) cells of a mating pair. The % fusion was calculated as the number of pairs fused over the total number of pairs counted multiplied by 100. It should be noted that at the earliest time points (≤ 1 h) there were fewer than 50 pairs available to count and so measurements of fusion at these times were less robust. Cell counts, as well as the fluorescence images shown in Figure 1, were obtained using a Zeiss Axio Imager M1 microscope equipped with an AxioCamMR3 camera. Prism software was used to generate scatter plots showing the kinetics of cellular pairing/fusion over time.

Western blotting analysis. *T. thermophila* cells modified to over-express HAP2 were induced to express a C-terminally HA-tagged HAP2 through the addition of 0.1 $\mu\text{g}/\text{mL}$ CdCl_2 to the 10 mM Tris starvation medium. Cell pellets (2×10^6 cells) were prepared every hour after induction for 10 h and 100 μL of $10 \times$ Roche cOmplete EDTA-free Protease Inhibitor Cocktail (Roche) was added prior to freezing. Frozen cell pellets were resuspended in 100 μL of 2x sodium dodecyl sulfate (SDS) loading buffer and boiled for 2 min before addition of DTT. Protein samples (equivalent to $\sim 1.5 \times 10^5$ cells) were separated on a 10% poly-acrylamide SDS gel in Tris-Glycine running buffer, and transferred to a 0.45 μm nitrocellulose membrane (Bio-Rad) for 1 h at 90 V. The blot in Figure S1D was blocked in a PBS-Tween containing 5% milk solution and probed with a 1/1000 dilution of anti-HA antibody (anti-HA.11, BioLegend, formerly Covance), followed by HRP-conjugated secondary anti-mouse IgG antibody at 1/1000 dilution (Southern Biotech). Signals were developed using a SuperSignal West Pico ECL kit and images acquired using a Syngene gel imager (Synoptics Ltd.).

For HAP2pp and Δenvpp conditions (Figure S4 B,C), co-transfected cells were lysed using with $1 \times$ radioimmunoprecipitation assay (RIPA) buffer (EMD Millipore) containing protease inhibitor cocktail (Roche). HAP2pp and Δenvpp pseudotyped particles were ultracentrifuged at 42,000 rpm for 2 h at 4°C, using a TLA-55 rotor with an Optima-MAX-E centrifuge (Beckman-Coulter). Viral pellets were resuspended in PBS. Lithium dodecyl sulfate (LDS) loading buffer and DTT were added to cell lysates and concentrated viral solutions, which were then heated at 95°C for 5 min. Protein samples were separated on a NOVEX Bis-Tris gel (Life Technologies) and transferred on a polyvinylidene fluoride membrane (GE Healthcare). Detection of HAP2-FLAG was performed using the mouse anti-FLAG tag antibody (FLAG M2 antibody, Sigma) and MLV capsid detection was performed using the mouse monoclonal anti-MLV capsid p30 (4B2, Abcam). Detection of Western blot signal was performed using an ECL kit (Pierce) and image acquisition was performed using an LAS-3000 imager (FujiFilm).

Template-based structural homology modeling. Initial protein homology modeling studies were conducted in 2013 using the Protein Homology/analogy Recognition Engine V2.0 (Phyre2) and the full-length amino acid sequence of *T. thermophila* HAP2 (GenBank: [KJ629172] as query. Phyre2 uses advanced remote homology detection algorithms involving alignments of amino acid sequences and predicted secondary structures to identify template hits within a hidden Markov model database of known structures. The identified top-scoring hits help form an alignment for the construction of a crude backbone structure while loop modeling and side chain placement are subsequently applied to finalize the predicted structure [S10]. Follow-up Phyre2 searches were carried out between 2014-2016 after other relevant class II structures were published (e.g. EFF-1), but yielded equivalent results. A Phyre batch processing query using 40 published [S2] HAP2 sequences was submitted in June 2015 (GenBank accession numbers of sequences listed in Table S1). Results yielded the top 20 ranked hits to known structures for each species' version of HAP2, as well as a confidence level for the prediction, a structural ID and template name, query start/stop sites of the aligned sequence, and a predicted structure of each ortholog based on the top ranking template hit. A partial list summarizing the top three ranking hits from both viral and non-viral template structures is shown in Table S1, as well as a list of the top viral hits in Figure 3E.

Submission of the *T. thermophila* HAP2 sequence to the RaptorX [S11,S12], CPHmodels3.0 [S13] and HHpred [S14] (Figure S2C) template-based homology modelers occurred between 2015-2016. Images of predicted structural models were generated using PyMol (Schrodinger, LLC). In all cases, domains I, II, and III were shaded red, yellow, and blue respectively according to the boundary locations for these domains inferred from the generated alignments of *T. thermophila* HAP2 and known class II fusion protein structures.

Expression of *T. thermophila* HAP2 in mammalian cell lines. A synthetic, codon-optimized version of the full length *T. thermophila* HAP2 gene was made (Genscript) with a 5'-*Hind*III restriction site and Kozak consensus sequence prior to the start codon, and a $3 \times$ FLAG epitope tag followed by a stop codon and an *Eco*RI restriction site at the 3'-end. For codon optimization of the *Tetrahymena thermophila* HAP2 gene, the amino acid sequence was first codon optimized for expression in *Homo sapiens* (GenScript) and was then submitted to the graphical codon usage analyzer (http://gcu.schoedl.de/sequential_v2.html) and manually altered (relative

adaptiveness values of <40%) to allow for optimal expression in both insect (*Drosophila melanogaster*) and human (*Homo sapiens*) cells. The codon-optimized HAP2 gene was digested at the aforementioned restriction sites, purified, and ligated to a similarly digested and purified pcDNA™3.1(+) mammalian expression vector (kindly provided by G. Whittaker). Resulting plasmids were verified by restriction analysis and sequencing (Cornell Biotechnology Resource Center). Purified plasmid DNA was concentrated by ethanol precipitation to ~1 µg/µL and used for transient transfections of mammalian cells.

Mammalian cell cultures of HEK-293T cells (ATCC), Huh-7 cells (Japan Health Science Research Resources Bank, Japan), Vero-E6 cells (ATCC), and BHK-21 cells (kindly provided by Mark Whitt) were maintained at 37°C 5% CO₂ in DMEM (Corning) supplemented with 10% fetal bovine serum (ThermoFisher), 10 mM HEPES (Corning), 100 IU/mL penicillin and 100 µg/mL streptomycin (Corning). For immunofluorescence analysis of transfected cells, 3-12.5 × 10⁴ Huh-7, Vero-E6, or BHK-21 cells were seeded in microscopy chamber slides (EMD Millipore) and incubated for 18 h at 37°C in a 5% CO₂ incubator. Cell supernatants were gently aspirated and replaced with 100 µL of warm Opti-MEM (ThermoFisher) media, then transfected by the further addition of 25 µL of an Opti-MEM-Lipofectamine 2000 (ThermoFisher) mixture containing the plasmid DNA encoding HAP2 or a pCAGGS empty vector control at a final concentration of 4 ng/µL, and incubated at 37°C in a 5% CO₂ incubator for 6 h. Transfection medium supernatant was then gently removed and replaced with 100 µL of warm DMEM supplemented with 10% fetal bovine serum and 10 mM HEPES but without penicillin/streptomycin and incubated 24 h at at 37°C in a 5% CO₂ incubator.

Pseudotyped particle production and infection assays. Murine Leukemia Virus (MLV)-based HAP2-pseudotyped particles (HAP2pp) were generated as previously described [S1]. 1 × 10⁶ HEK-293T cells were seeded in six-well plates and incubated at 37°C in a 5% CO₂ incubator for 18 h. Cells were co-transfected with pcDNA-HAP2-FLAG plasmid (HAP2pp), or pCAGGS empty vector control (Δenvpp), or VSV-G encoding plasmid (VSV-Gpp), along with MLV Gag-Pol packaging construct and the MLV transfer vector (encoding a luciferase reporter gene), using Lipofectamine 2000 transfection reagent (Life Technologies) according to manufacturer's instructions. The cells were incubated at 37°C in a 5% CO₂ incubator for 48 h. Supernatants containing released pseudotyped particles were harvested and filtered through 0.45 µm membranes and stored at -80°C.

For infection assays, 2.5 × 10⁵ HEK-293T, or Huh-7, or Vero-E6 cells were seeded in 24-well plates and incubated at 37°C in a 5% CO₂ incubator for 18 h. Cells were transfected with either pCAGGS empty-vector control or pcDNA-HAP2-FLAG plasmid and incubated at 37°C for 24 h at 37°C in a 5% CO₂ incubator. The cells were washed with PBS, and 200 µL of pseudotyped particles were added to cells and incubated at 37°C in a 5% CO₂ incubator for 2 h. Complete medium was then added and cells were incubated at 37°C in a 5% CO₂ incubator for 72 h, after which luciferase activity was measured using Luciferase Assay Kit (Promega), and luminometer readings performed with a Glomax 20/20 system (Promega). Experimental values were plotted using Prism 7 (GraphPad) and are average relative luciferase units of three replicates (n = 3) with error bars representing standard deviation (s.d.).

Lipids and peptides used for biophysical studies. The lipids POPC, POPG, the chain spin label, 5PC, and the head group spin label, dipalmitoylphosphatidyl-tempo-choline (DPPTC), were purchased from Avanti[®] Polar Lipids. Cholesterol was purchased from Sigma. The wild type HAP2 and Dengue virus fusion peptides were synthesized by ChinaPeptides Co., Ltd., and the Influenza virus fusion peptides by SynBioSci Co. Sequences of HAP2 negative control peptides were randomized using Shuffle Protein [S15] and these randomized peptides were synthesized by New England Peptide Inc. All sequences of experimental and control peptides had the solubility tag "GGGKKKK" added to their C' terminal ends [S16–S18] and are shown in Figure 5A. The structures of the spin labeled lipids used in these experiments are shown in Figure 5C,D.

Preparation of membrane vesicles. The volumes of POPC, POPG, cholesterol and 0.5% (mol:mol) spin-labeled lipids in chloroform were mixed according to a 5:2:3 vol:vol ratio of POPC:POPG:Chol and dried by N₂ flow. The mixture was evacuated in a vacuum drier overnight to remove any trace of chloroform. To prepare multilamellar vesicles (MLVs), the lipids were resuspended and fully hydrated using 1 mL of pH 5 buffer (5 mM HEPES, 10 mM MES, 150 mM NaCl, pH 5) at room temperature for 2 h. To prepare small unilamellar vesicles (SUVs), the lipids were resuspended in pH 5 buffer and sonicated in ice bath for 20 min. To prepare large unilamellar vesicles (LUVs), the lipids were frozen and thawed 5 times before they were extruded in an Avanti extruder through a membrane with 100 nm pore size.

Electron spin resonance (ESR) spectroscopy and nonlinear least-squares fit of ESR spectra. To prepare the samples for lipid ESR study, a stock solution of the Fusion Peptide (FP) (1 mg/mL) was added to the

lipid POPC:POPG:Chol=5:2:3 MLV dispersion (above) at the experimentally indicated ratios. After 20 min of incubation, the dispersion was centrifuged at 13,000 rpm for 10 min. The pellet was transferred to a quartz capillary tube for ESR measurement. ESR spectra were collected on an ELEXSYS ESR spectrometer (Bruker Instruments) at X-band (9.5 GHz) at 25°C using a N2 Temperature Controller (Bruker Instruments). The ESR spectra from the labeled lipids were analyzed using the NLLS fitting program based on the stochastic Liouville equation [S19] using the MOMD (Microscopic Order Macroscopic Disorder) model as in previous studies [S20–S23]. The fitting strategy is the same as previously reported [S24]. S_0 is defined as follows: $S_0 = \langle D_{2,00} \rangle = \langle 1/2(3\cos^2\theta - 1) \rangle$, where $D_{2,00}$ is the Wigner rotation matrix elements and θ is the polar angle for the orientation of the rotating axes of the nitroxide bonded to the lipid relative to the director of the bilayer, i.e. the preferential orientation of lipid molecules [S21,S25], and the angular brackets imply ensemble averaging. S_0 indicates how well the chain segment to which the nitroxide is attached, is aligned along the normal to the lipid bilayer [S24].

Circular dichroism spectroscopy (CD). Fusion peptides (0.2 mg/mL in pH 5 solution) were mixed with SUVs composed of POPC:POPG:Chol=5:2:3 at a ratio of 1:100 peptide:lipid at room temperature for 10 min before measurement. The CD spectra were collected at 25°C on an AVIV Model 202-01 Circular Dichroism Spectrometer (AVIV biomedical Inc.). The signals from pure SUVs or pure solution were subtracted from the sample spectra as blanks. The CD spectra were analyzed using K2D3[S26].

Fluorescence dequenching assays. The protocol for fluorescence dequenching assays to monitor vesicle fusion was adopted from a previous study [S27]. Fluorescently labeled LUVs (2.5 μ M, final concentration) containing 2% Octadecyl Rhodamine B chloride (R18, Molecular Probe, ThermoFisher Scientific) and unlabeled LUV (22.5 μ M, final concentration) were mixed in 1 mL of pH 5 buffer. Fusion peptides were then added from concentrated stock solutions to give a 1 μ M final concentration of each peptide. 10% Triton X-100 was added to achieve a 1% final concentration after fusion reactions were complete. The fluorescence spectra were collected on a Varian Cary Eclipse Fluorescence Spectrometer (Agilent Technologies). Fluorescence intensities of the samples before addition of fusion peptides and after the addition of Triton X-100 were used to set the baseline (0%) and 100% fusion levels, respectively. The fluorescence yields of the experimental samples were normalized to these levels to determine % lipid mixing [S28]. Fluorescence intensity variations due to volume changes were corrected in each case. All experiments were performed at least 3 times and representative curves are shown.

Statistical information. All statistical tests were performed using Prism 7 software (GraphPad Inc.). Sample sizes for the functional analyses of HAP2 mutant strains by flow cytometry (Figure 2G, Figure 4B,D, and Figure S1) are listed in parentheses below with the total number of biological replicates over all experiments (that is, total number of individual matings performed) listed first, followed by the total number of independent experiments performed for each cross: WT ^{α} x WT ^{β} (42, 17); Δ HAP2 ^{α} x WT ^{β} (20, 8); WT ^{β} x Δ HAP2 ^{α} (31, 10); WT ^{α} x Δ HAP2 ^{β} (26, 10); Genomic ^{β} x WT ^{α} (4, 2); Genomic ^{β} x Δ HAP2 ^{α} (5, 3); cDNA ^{β} x WT ^{α} (5, 4); Δ HAP2 domain (6, 5); Δ DENV region (5, 4); Δ Fusion Loop (10, 5); DENV FL Rescue (12, 3); Δ 510-513 (7, 3); HAP2 FL Rescue (6, 3); FQY131-3AAA (8, 3); CC147-8SS (16, 4); R164A (9, 3); LNL171-3AAA (9, 3); C₅→S (4, 3); C₈→S (13, 5); Δ Basic Domain (10, 3); Δ C' term (9, 3); O.E. HAP2 x Δ HAP2 ^{α} (17, 4); O.E. HAP2 x WT ^{α} (13, 3). No pre-determined power analyses were performed for calculating sample sizes. The criterion for including any given cross in cell fusion analysis was a pairing frequency of >60% in mating cultures 3 h post mixing. No randomization was applied to samples and the investigators were not blinded to sample identity during the experiments.

A one-sided non-parametric Kruskal Wallis test with Dunn's multiple comparisons post-test was applied to the cellular fusion data because not all data sets passed the D'Agostino-Pearson Omnibus normality test. Associated P values and standard deviations (as measurements of variance) are shown in the text, figures, and figure legends. From these tests, there was a significant difference found in the percent fusion for data from the 8 different crosses in Figure 2G ($H_7 = 116.3$, $P < 0.0001$). For Figure 4C, a significant difference was also found in comparisons to either data from WT ^{α} x WT ^{β} or WT ^{α} x KO ^{β} crosses respectively ($H_{11} = 123.2$ and $H_{10} = 85.78$, and $P < 0.0001$). Similarly, when the groups shown in Figure 4D were compared to data from WT ^{α} x WT ^{β} and WT ^{α} x KO ^{β} crosses respectively, significant differences were also found ($H_4 = 14.92$, $P = 0.0049$ and $H_4 = 48.17$, and $P < 0.0001$).

Data availability. All relevant data are available from the authors. Specifically, HAP2 mutant cell lines have been deposited in the *Tetrahymena* Stock Center at Cornell University and are available for use by the community at-large. GenBank accession numbers to relevant sequences used in this study are shown both below and in Table S1. The entire Phyre batch processing data set in its raw form along with the corresponding Phyre2-

predicted structures and homology models from other template-based prediction tools will be made available upon request to the corresponding author.

Accession numbers. The accession number for the *T. thermophila* HAP2 protein sequence is GenBank: [KJ629172]. Additional information on the gene is available through the Tetrahymena Genome Database (TGD) Wiki (<http://ciliate.org/index.php/home/welcome>) under the accession number: TGD: [TTHERM_01075640].

***Tetrahymena thermophila* strains.**

Strain^a	Genotype (micronucleus)	Genotype; Phenotype (macronucleus)
CU428.2	<i>CHX1/CHX1; mpr1-1/mpr1-1</i>	<i>MPR1, CHX1; mp-s, cy-s, VII</i>
CU427.4	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1; mp-s, cy-s, VI</i>
ΔHAP2-428	<i>CHX1/CHX1; mpr1-1/mpr1-1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4]; mp-s, cy-s, pm-r, VII</i>
ΔHAP2-427	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4]; mp-s, cy-s, pm-r, VI</i>
HAP2cDNAResc427	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2cDNA(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
HAP2genomicResc427	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-3(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔHAP2 Domain	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N2(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔDENV Region	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N3(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔFusion Loop	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N5(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔDENV FL Rescue	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N4(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
Δ510-513	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N6(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔHAP2 FL Rescue	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N5R(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔFQY131-3AAA	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N71(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔCC-147-8SS	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N72(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔR164A	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N73(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔLNL171-173AAA	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-N74(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔC ₅ →S	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-C3(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔC ₈ →S	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-C5(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔBasic Domain	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-C2(3'cy2)]; mp-s, pm-r, cy-r, VI</i>
ΔC' term	<i>chx1-1/chx1-1; MPR1/MPR1</i>	<i>MPR1, CHX1, hap2-1[Δ::neo4/hap2-C1(3'cy2)]; mp-s, pm-r, cy-r, VI</i>

^aCU428.2 and CU427.4 are functional heterokaryons that are phenotypically sensitive (s) to 6-methylpurine (mp) and cycloheximide (cy) respectively, due to markers present in their macronuclei, but homozygous for resistance (r)

N6HyHAP2rev	CACTGCAGGATTAGGATGACTTTGATTATTACTA GCCTGTGTTATTTCC	Δ510-513
N7_R164A_F	TGATCTATCAGCTGGTAAAGTGTGCTATGC	R164A
N7_R164A_R	TTACCCATGCCTAATATATC	R164A
N7_CC147SS_F	CTAAGGTTATAGTAGTTATTGCTCTCTATCAGAT ATATTAG	CC147-8SS
N7_CC147SS_R	CTGTCCAAAATCTTTTAGCC	CC147-8SS
N7_FQY131AAA_F	AGCTGACTCTAAAAGGCTAAAAGATTTTG	FQY131-3AAA
N7_FQY131AAA_R	GCAGCTTTACAAGTTGGACTACTATC	FQY131-3AAA
N7_LNL171AAA_F	AGCTGGTGCAGGATCAGCAACA	LNL171-3AAA
N7_LNL171AAA_R	GCAGCAGCATAGCACACTTTACCTC	LNL171-3AAA
HAP2 5' FlankFor	GTTATTTTCAGCATCTTCTTTCATTTG	Genotyping
HAP2 3' FLANK REV	ATCTCTTCTGATCATAGAGCACC	Genotyping
RV3rev	GACATTAAAGCAAGTTAAGCATAAATAAAG	All Truncations
Mtt1RevSeq	AATACGAAACTGATTTTATGCAA	Genotyping
PreStagsForKpnI	GTAATCggtaccGAAGTTTTGTTCCAAGGTCCC	Over-expression HAP2
HistagRevSacI	GATTACgagctcTCAGTGGTGGTGGTGGTGGT	Over-expression HAP2

^aAll primers used in this project were supplied by Sigma-Aldrich. Primers were stored as 100 μM stock solutions, and diluted to 10 μM working solutions prior to use in PCR reactions. Melting temperatures were determined using Modified Breslauer's thermodynamics, dH and dS parameters as recommended by the manufacturer. Added restriction sites are shown in lower case letters.

^bAll constructs made with these primers were subjected to Sanger sequencing to verify accuracy.

Supplemental References

- S1. Bartosch, B., Dubuisson, J., and Cosset, F.-L. (2003). Infectious Hepatitis C Virus Pseudo-particles Containing Functional E1–E2 Envelope Protein Complexes. *J. Exp. Med.* *197*, 633–642.
- S2. Liu, Y., Pei, J., Grishin, N., and Snell, W.J. (2015). The cytoplasmic domain of the gamete membrane fusion protein HAP2 targets the protein to the fusion site in *Chlamydomonas* and regulates the fusion reaction. *Development* *142*, 962–971.
- S3. Bruns, Peter J, and Brussard, Trudy B. (1974). Positive selection for mating with functional heterokaryons in *Tetrahymena pyriformis*. *Genetics* *78*, 831–841.
- S4. Cole, E.S., Cassidy-Hanley, D., Pinello, J.F., Zeng, H., Hsueh, M., Kolbin, D., Ozzello, C., Jr, T.G., Winey, M., and Clark, T.G. (2014). Function of the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate. *Curr. Biol. CB* *24*, 2168–2173.
- S5. Cassidy-Hanley, D., Bowen, J., Lee, J.H., Cole, E., VerPlank, L.A., Gaertig, J., Gorovsky, M.A., and Bruns, P.J. (1997). Germline and somatic transformation of mating *Tetrahymena thermophila* by particle bombardment. *Genetics* *146*, 135–147.
- S6. Shang, Y., Song, X., Bowen, J., Corstanje, R., Gao, Y., Gaertig, J., and Gorovsky, M.A. (2002). A robust inducible-repressible promoter greatly facilitates gene knockouts, conditional expression, and overexpression of homologous and heterologous genes in *Tetrahymena thermophila*. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 3734–3739.
- S7. Cassidy-Hanley, D., Smith, H.R., and Bruns, P.J. (1995). A simple, efficient technique for freezing *Tetrahymena thermophila*. *J. Eukaryot. Microbiol.* *42*, 510–515.
- S8. Wolfe, J. (1976). G1 arrest and the division/conjugation decision in *Tetrahymena*. *Dev. Biol.* *54*, 116–126.
- S9. Parish, C.R. (1999). Fluorescent dyes for lymphocyte migration and proliferation studies. *Immunol. Cell Biol.* *77*, 499–508.
- S10. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* *10*, 845–858.
- S11. Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., and Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* *7*, 1511–1522.
- S12. Peng, J., and Xu, J. (2011). Raptorx: Exploiting structure information for protein alignment by statistical inference. *Proteins Struct. Funct. Bioinforma.* *79*, 161–171.
- S13. Nielsen, M., Lundegaard, C., Lund, O., and Petersen, T.N. (2010). CPHmodels-3.0—remote homology modeling using structure-guided sequence profiles. *Nucleic Acids Res.* *38*, W576–W581.
- S14. Söding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* *33*, W244–248.
- S15. Stothard, P. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques* *28*, 1102, 1104.
- S16. Lai, A.L., and Freed, J.H. (2014). HIV gp41 fusion peptide increases membrane ordering in a cholesterol-dependent fashion. *Biophys. J.* *106*, 172–181.

- S17. Han, X., and Tamm, L.K. (2000). A host-guest system to study structure-function relationships of membrane fusion peptides. *Proc. Natl. Acad. Sci. U. S. A.* *97*, 13097–13102.
- S18. Han, X., Bushweller, J.H., Cafiso, D.S., and Tamm, L.K. (2001). Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nat. Struct. Biol.* *8*, 715–720.
- S19. Budil, David E., Lee, Sanghyuk, Saxena, Sunil, and Freed, Jack H. (1996). Nonlinear-Least-Squares Analysis of Slow Motion EPR Spectra in One and Two Dimensions Using a Modified Levenberg-Marquardt Algorithm. *J. Magn. Reson. Series A*, 155–189.
- S20. Ge, M., and Freed, J.H. (2003). Hydration, Structure, and Molecular Interactions in the Headgroup Region of Dioleoylphosphatidylcholine Bilayers: An Electron Spin Resonance Study. *Biophys. J.* *85*, 4023–4040.
- S21. Ge, M., and Freed, J.H. (2009). Fusion Peptide from Influenza Hemagglutinin Increases Membrane Surface Order: An Electron-Spin Resonance Study. *Biophys. J.* *96*, 4925–4934.
- S22. Ge, M.T., Costa-Filho, A., Gidwani, A., Holowka, D., Baird, B., and Freed, J. (2001). The structure of bleb membranes of RBL-2H3 cell is heterogeneous: An ESR study. *Biophys. J.* *85*, 4023–4040.
- S23. Smith, A.K., and Freed, J.H. (2009). Determination of tie-line fields for coexisting lipid phases: an ESR study. *J. Phys. Chem. B* *113*, 3957–3971.
- S24. Lai, A.L., and Freed, J.H. (2015). The Interaction between Influenza HA Fusion Peptide and Transmembrane Domain Affects Membrane Structure. *Biophys. J.* *109*, 2523–2536.
- S25. Liang, Z., and Freed, J.H. (1999). An Assessment of the Applicability of Multifrequency ESR to Study the Complex Dynamics of Biomolecules. *J. Phys. Chem. B* *103*, 6384–6396.
- S26. Louis-Jeune, C., Andrade-Navarro, M.A., and Perez-Iratxeta, C. (2012). Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins Struct. Funct. Bioinforma.* *80*, 374–381.
- S27. Ravoo, B.J., Weringa, W.D., and Engberts, J.B. (1999). Membrane fusion in vesicles of oligomerizable lipids. *Biophys. J.* *76*, 374–386.
- S28. Lai, A.L., Moorthy, A.E., Li, Y., and Tamm, L.K. (2012). Fusion Activity of HIV gp41 Fusion Domain Is Related to Its Secondary Structure and Depth of Membrane Insertion in a Cholesterol-Dependent Fashion. *J. Mol. Biol.* *418*, 3–15.