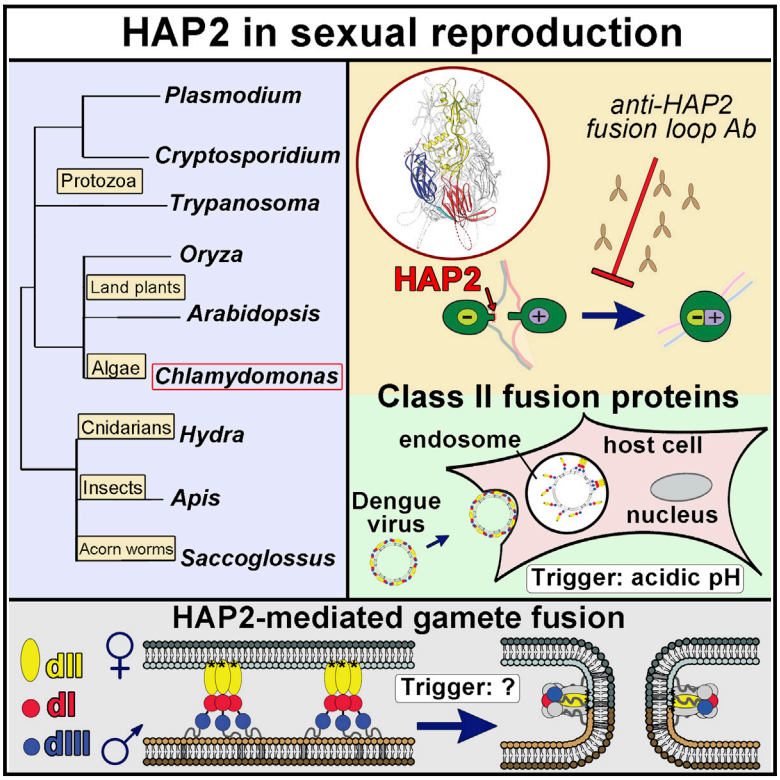
Article

Cell

The Ancient Gamete Fusogen HAP2 Is a Eukaryotic Class II Fusion Protein

Graphical Abstract



Highlights

- The primordial gamete fusogen HAP2 exhibits homology to class II viral fusion proteins
- HAP2 inserts into the target gamete membrane via a hydrophobic fusion loop
- HAP2 links virus entry into target cells and the origins of sexual reproduction
- HAP2 is a sex-specific target for blocking fertilization in multiple kingdoms

Authors

Juliette Fédry, Yanjie Liu, Gérard Péhau-Arnaudet, ..., William J. Snell, Félix A. Rey, Thomas Krey

Correspondence

wsnell1@umd.edu (W.J.S.), felix.rey@pasteur.fr (F.A.R.), krey.thomas@mh-hannover.de (T.K.)

In Brief

Gamete fusion across eukaryotic branches uses an ancient factor homologous to viral fusion proteins.

Fédry et al., 2017, Cell *168*, 904–915
February 23, 2017
http://dx.doi.org/10.1016/j.cell.2017.01.024

CellPress

Cell

Article

The Ancient Gamete Fusogen HAP2 Is a Eukaryotic Class II Fusion Protein

Juliette Fédry,[1,2,3,4,11] Yanjie Liu,[5,11] Gérard Péhau-Arnaudet,[6] Jimin Pei,[7] Wenhao Li,[5] M. Alejandra Tortorici,[1,2] François Traincard,[8] Annalisa Meola,[1,2] Gérard Bricogne,[9] Nick V. Grishin,[7] William J. Snell,[5,12,*] Félix A. Rey,[1,2,13,*] and Thomas Krey[1,2,4,10,*]

[1]Unité de Virologie Structurale, Institut Pasteur, 25-28 Rue du Docteur Roux, 75724 Paris, France
[2]CNRS UMR 3569, 25-28 Rue du Docteur Roux, 75724 Paris, France
[3]Université Paris Descartes Sorbonne Paris Cité, Institut Pasteur, Rue du Docteur Roux, 75015 Paris, France
[4]Institute of Virology, Hannover Medical School, 30625 Hannover, Germany
[5]Department of Cell Biology, University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard, Dallas, TX 75390-9039, USA
[6]Ultrapole, UMR 3528, CNRS, Institut Pasteur, 25-28 rue du Docteur Roux, 75015 Paris, France
[7]Departments of Biophysics and Biochemistry and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard, Dallas, TX 75390-9039, USA
[8]Plateforme Ingenierie des Anticorps, Department Biologie Structurale et Chimie, Institut Pasteur, 75015 Paris, France
[9]Global Phasing Ltd., Sheraton House, Castle Park, Cambridge CB3 0AX, UK
[10]German Center for Infection Research (DZIF), 30625 Hannover, Germany
[11]Co-first author
[12]Present address: Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742, USA
[13]Lead Contact
*Correspondence: wsnell1@umd.edu (W.J.S.), felix.rey@pasteur.fr (F.A.R.), krey.thomas@mh-hannover.de (T.K.)
http://dx.doi.org/10.1016/j.cell.2017.01.024

SUMMARY

Sexual reproduction is almost universal in eukaryotic life and involves the fusion of male and female haploid gametes into a diploid cell. The sperm-restricted single-pass transmembrane protein HAP2-GCS1 has been postulated to function in membrane merger. Its presence in the major eukaryotic taxa—animals, plants, and protists (including important human pathogens like *Plasmodium*)—suggests that many eukaryotic organisms share a common gamete fusion mechanism. Here, we report combined bioinformatic, biochemical, mutational, and X-ray crystallographic studies on the unicellular alga *Chlamydomonas reinhardtii* HAP2 that reveal homology to class II viral membrane fusion proteins. We further show that targeting the segment corresponding to the fusion loop by mutagenesis or by antibodies blocks gamete fusion. These results demonstrate that HAP2 is the gamete fusogen and suggest a mechanism of action akin to viral fusion, indicating a way to block *Plasmodium* transmission and highlighting the impact of virus-cell genetic exchanges on the evolution of eukaryotic life.

INTRODUCTION

Fusion of haploid cells to form a diploid zygote is the defining event of sexual reproduction in eukaryotes (Lillie, 1913). In organisms from every eukaryotic taxon, the plasma membranes of gametes of opposite sex or mating type come into intimate contact and then fuse to form the zygote (Bianchi et al., 2014; Okabe, 2013; Sinden et al., 1976; Snell and Goodenough, 2009; Sprunck et al., 2012). In spite of the importance of gamete fusion, very little is known about the molecular mechanisms of the membrane fusion reaction between gametes, and a bona fide fusion protein has not been formally identified. The best candidate to date is the ancient gamete plasma membrane protein HAP2, whose presence in green algae, higher plants, unicellular protozoa, cnidarians, hemichordates, and arthropods (Cole et al., 2014; Ebchuqin et al., 2014; Johnson et al., 2004; Kawai-Toyooka et al., 2014; Liu et al., 2008; Mori et al., 2006; Steele and Dana, 2009) indicates it was likely present in the last eukaryotic common ancestor (LECA) (Wong and Johnson, 2010). HAP2 was first identified in a screen for male sterility in the flowering plant *Arabidopsis thaliana* (Johnson et al., 2004) and later—under the name GCS1 (Mori et al., 2006)—as a sperm-specific protein shown to be required at an unidentified step in sperm-egg fusion (Mori et al., 2006; von Besser et al., 2006). A screen for genes essential for gamete fusion in the green alga *Chlamydomonas* independently uncovered HAP2, showing that it is expressed only in *minus* gametes and is exclusively present on an apically localized membrane protuberance termed the *minus* mating structure (Liu et al., 2008) (see Figure 1A for a diagram of *Chlamydomonas* fertilization). Studies in *Chlamydomonas* and *Plasmodium* (the pathogen causing malaria in humans) revealed that *hap2* mutant gametes in both organisms were fully capable of robust adhesion to gametes of the opposite mating type or sex, but merger of the lipid bilayers was abrogated (Liu et al., 2008). In both organisms, adhesion relies on proteins that are species-limited: FUS1 in *Chlamydomonas plus* gametes and its unidentified receptor in *minus* gametes (Misamore et al., 2003), and p48/45 in *Plasmodium berghei* gametes (van Dijk et al., 2001).

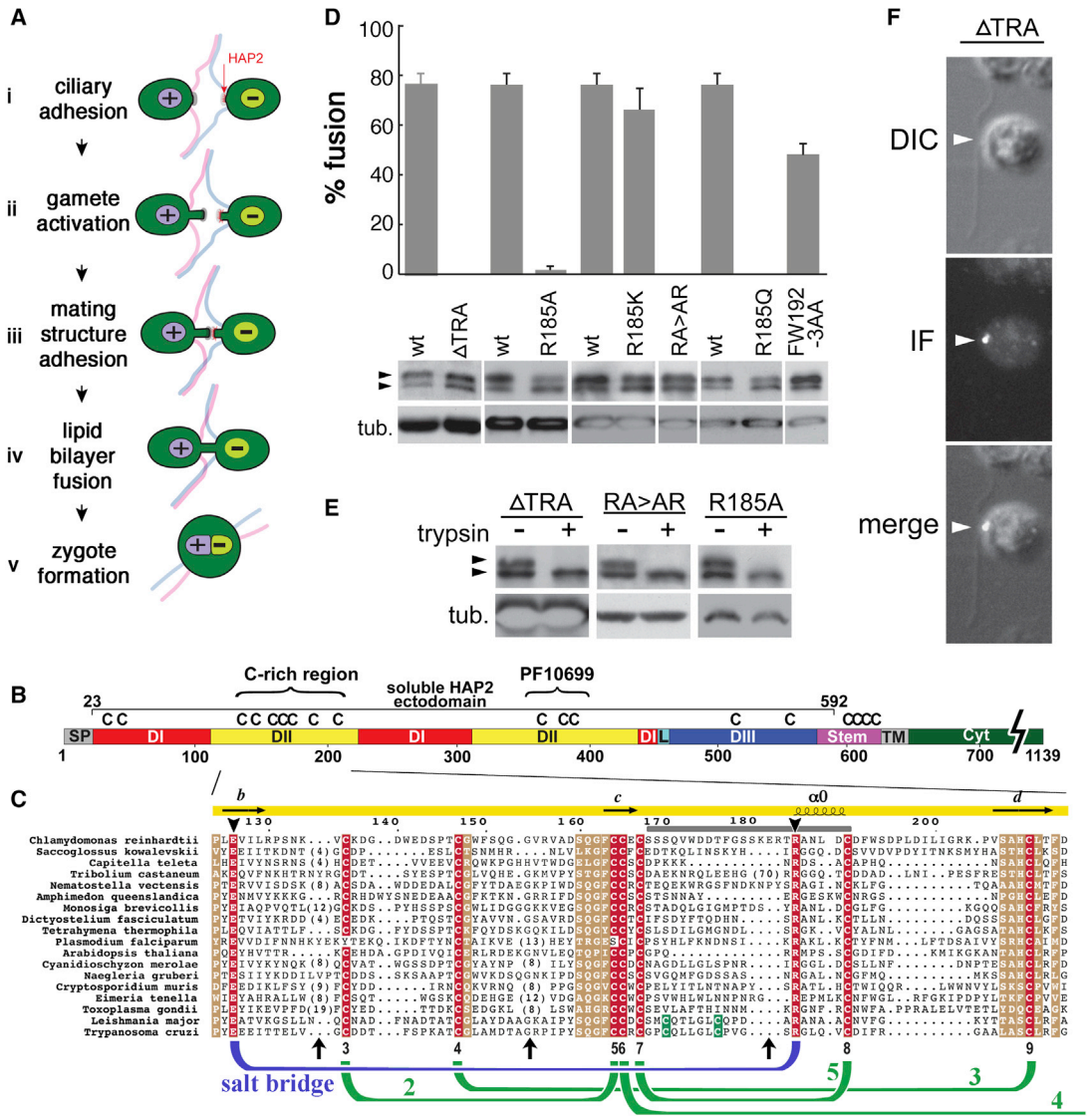


Figure 1. Identification of a HAP2 Region Containing Elements Essential for HAP2 Function in Gamete Fusion

(A) Schematic representation of *Chlamydomonas* gamete fusion. Adhesion through mating-type-specific ciliary adhesion proteins brings gametes together (i) and activates erection of mating structures on each gamete (ii), resulting in adhesion of the mating structure tips through separate cell-specific adhesion proteins (iii). Within seconds HAP2 (indicated and labeled in red with a vertical arrow) induces lipid merger to complete the membrane fusion reaction (iv), followed by coalescence of the two gametes into a zygote (v).

(B) Primary structure of *Chlamydomonas* HAP2 showing the PFAM domain PF10699, conserved cysteine residues and the cysteine-rich region discussed in the text. SP, signal peptide; TM, transmembrane segment; Cyt, cytosolic tail. Domains DI, DII, DIII, the domain I-III linker (L), and the Stem are presented in Figure 5.

(C) Alignment of the cysteine-rich region of HAP2 proteins from representative organisms. Indicated above the alignment are the secondary structure elements taken from the *C. reinhardtii* HAP2 crystal structure (presented in Figures 4 and 5). Conserved and semi-conserved residues are in white font on a red or a beige background, respectively. Cysteine residues are numbered sequentially below the alignment and their disulfide connectivity is drawn in green with the disulfide bonds numbered as in Figure 5. The conserved salt bridge between R_{185} and E_{126} (arrowheads above the sequence) is drawn in blue below the sequences (see Figures 5B and 5C). A gray bar above the alignment shows the 168-190 peptide used to immunize rabbits (see also Figure S1). Black arrows below the alignment point to positions where HAP2 proteins from other species carry insertions, with their length given within the corresponding sequence in parenthesis.

(D) Gamete fusion activity and protein expression of HAP2 mutants. The top panel summarizes the fusion assays, which were performed at least as duplicates with two counts on fusion rate in each experiment and results are shown as mean ± SD. The immunoblots in the lower panel show that the various mutants were expressed in the gametes, with arrowheads pointing to the HAP2-HA doublet bands.

(E) Mutant HAP2 is exposed at the gamete surface. The sensitivity of the upper form of HAP2 to trypsin indicates that it is accessible at the cell surface. Tubulin staining (tub.) indicates equivalent loadings in the various gels.

(F) Differential interference contrast (DIC) microscopy and HA immunofluorescence demonstrate the expected localization of HAP2-ΔTRA at the site of the *minus* mating structure (arrowheads) between the two flagella.

See also Figure S1.

Based on these findings, which have since been confirmed in *Arabidopsis thaliana* (Mori et al., 2014) and the ciliated protozoan *Tetrahymena thermophila* (Cole et al., 2014), a model emerged positing that HAP2, a single-pass transmembrane protein, functions after species-limited adhesion in the membrane fusion process between gametes (Liu et al., 2008). Furthermore, in all of these organisms, HAP2 is required in only one of the two gametes, raising the possibility that it may function similarly to fusion proteins of enveloped viruses (Wong and Johnson, 2010; Harrison, 2015).

To understand the function of HAP2 at the molecular level we carried out concerted bioinformatic, functional, and X-ray structural analyses of HAP2 from *Chlamydomonas reinhardtii*. Initial bioinformatic analyses identified weak similarity to class II fusion proteins, revealing a segment within a cysteine-rich portion of HAP2 that could potentially correspond to the fusion loop. We demonstrate by mutational analysis and fusion-blocking antibodies targeting this segment that it has elements that are essential for HAP2 function. Finally, we show that the recombinant HAP2 ectodomain is monomeric, but inserts into liposomes by concomitantly forming trimers, the X-ray structure of which revealed a class II fusion protein fold in the typical trimeric post-fusion "hairpin" conformation.

RESULTS AND DISCUSSION

Bioinformatic and Functional Analyses of HAP2

HAP2 has 16 conserved cysteine residues with a signature distribution in the ectodomain (Figure 1B). Early alignments of HAP2 family members identified a characteristic ~50 aa domain (residues 352–399 in *Chlamydomonas* HAP2) with several conserved residues that was designated the HAP2/GCS1 PFAM domain (PF10699) (see http://pfam.xfam.org/family/pf10699) (Finn et al., 2016). A previous mutagenesis analysis in *Chlamydomonas* failed to identify functional properties in the PF10699 domain, as the mutant proteins tested either were not transported to the mating structure, or were nearly indistinguishable from wild-type in their ability to support fusion with *plus* gametes (Liu et al., 2015). Database searches for additional conserved regions using the HHpred protein homology detection server (Söding et al., 2005) indicated that a cysteine-rich region in the N-terminal half of the ectodomain exhibited weak similarity to class II fusion proteins. In particular, HHpred identified a polypeptide segment in *C. reinhardtii* HAP2 (170–204, SQVWDDTFGSSKERT**R**ANLD**C**DFWSDPLDILIGRK) that fell in the fusion loop region of the flavivirus envelope protein E in the resultant amino-acid-sequence alignment (Figure S1). Analysis of HAP2 orthologs showed that the sequence in this region is highly variable, with a number of deletions and insertions and is framed at each side by relatively conserved segments: amino acids (aa) 159–167 upstream (including conserved cysteines 5–7) and aa 208–219 downstream (including conserved cysteine 9) (Figures 1B and 1C). Only amino acids R_{185} and C_{190} (in bold in the sequence above) within the identified segment are conserved, suggesting that they may play a role in HAP2 function.

To investigate the functional importance of this segment we transformed *Chlamydomonas* wild-type (WT) or mutant *HAP2* transgenes carrying an influenza virus hemagglutinin tag (HA) into a fusion-defective, *hap2* mutant strain (Liu et al., 2008) and assessed HAP2-HA expression and trafficking to the mating structure as well as fusion of the transformed *hap2 minus* gametes with WT *plus* gametes. HAP2-HA was detected in *hap2 minus* gametes transformed with WT *HAP2-HA* as the expected doublet in SDS-PAGE/immunoblotting (Figure 1D), the upper form of which was present on the cell surface as assessed by its sensitivity to protease treatment of live gametes (Liu et al., 2008). All mutant proteins were expressed at levels similar to wild-type HAP2-HA (Figure 1D), trafficked to the cell surface as assessed by their sensitivity to trypsin treatment of live gametes (examples shown in Figure 1E), and localized at the mating structure (example shown in Figure 1F). Thus, any defects in gamete fusion could be ascribed directly to the functional properties of the mutant HAP2 proteins. HAP2 with a deletion of residues $_{184}$TRA$_{186}$ (HAP2-ΔTRA-HA), which includes the conserved R_{185} was non-functional and failed to rescue fusion in the *hap2* mutant when mixed with wild-type *plus* gametes (Figure 1D). A mutant HAP2 with a lysine sustituted for the conserved R_{185} (the HAP2-R_{185}K-HA mutant) was fully functional, whereas the HAP2-R_{185}A-HA or HAP2-R_{185}Q-HA mutants failed to rescue fusion (Figure 1D). A reverse-order HAP2-RA$_{185\text{-}86}$ mutant (HAP2-R_{185}A-A_{186}R-HA) also was non-functional, indicating that a positively charged residue at position 185 is essential for the fusion activity. Finally, *hap2* minus gametes expressing HAP2-F_{192}A-W_{193}A-HA were impaired in fusion, although fusion was not abolished by these mutations, indicating that these nearby aromatic residues also play a role in HAP2 fusion function (Figure 1D). Thus, the segment HAP2$_{170\text{-}204}$, bounded by a pair of conserved cysteines, contains residues that are dispensable for protein expression, folding and localization, but are essential for the membrane fusion activity.

An Antibody Targeting the HHpred-Identified Region Blocks Gamete Fusion

In an independent approach to examine the function of the HHpred-identified region, we generated a rabbit antibody against a synthetic peptide, HAP2$_{168\text{-}190}$, spanning the functionally important R_{185} residue. The affinity-purified antibody (α-HAP2$_{168\text{-}190}$) immunoprecipitated epitope-tagged HAP2-HA from lysates of *HAP2-HA minus* gametes (Figure 2A), confirming its reactivity with HAP2. To test whether α-HAP2$_{168\text{-}190}$ interfered with gamete fusion, we incubated *minus* gametes with undiluted antibody, mixed them with *plus* gametes, and determined the percentage of gametes that had fused to form zygotes. Pre-incubation of *minus* gametes with α-HAP2$_{168\text{-}190}$ had no effect on motility or adhesion but inhibited gamete fusion by over 75%, whereas pre-incubation with a control IgG had no effect on fusion (Figures 2B and 2C). Antibody dilution resulted in a loss of fusion-blocking activity, suggesting a low concentration of HAP2-specific antibodies in the polyclonal mixture, probably due to a low immunogenicity of the synthetic peptide. Pre-incubation of *plus* gametes with the antibody did not affect their ability to fuse; and the fusion-blocking activity of α-HAP2$_{168\text{-}190}$ was neutralized by pre-incubation with the HAP2$_{168\text{-}190}$ peptide, but not with a control peptide (Figure 2C), further documenting the specificity of the antibody. Finally,

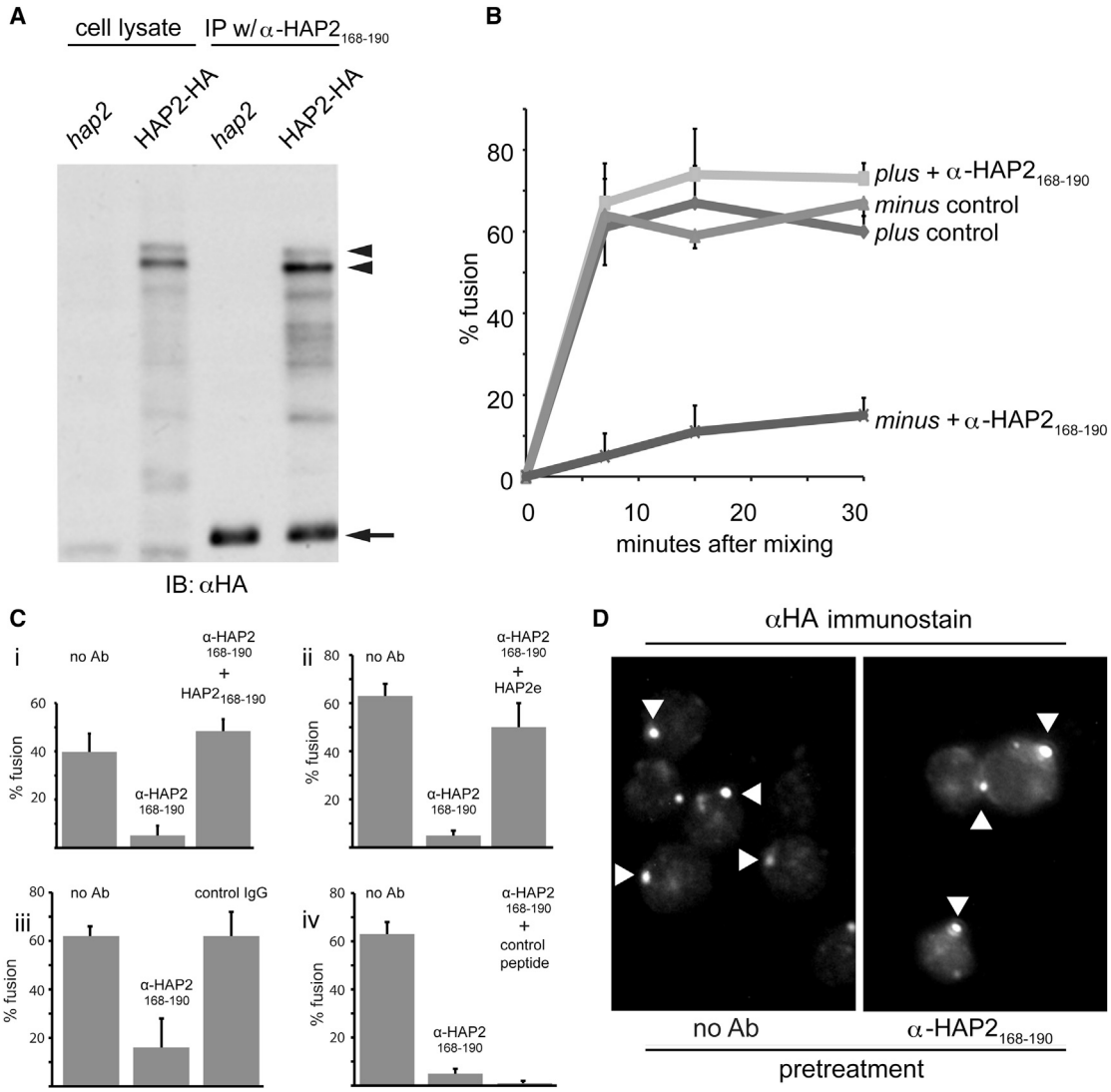


Figure 2. An Antibody Targeting a HAP2 S_{168}-C_{190} Peptide Inhibits Gamete Fusion

(A) Anti-HAP2$_{168-190}$ peptide antibody (α-HAP2$_{168-190}$) immunoprecipitates HAP2-HA (arrowheads) from *hap2/HAP2-HA* gametes, but not *hap2* gametes. The immunoblot was revealed with an α-HA Mab. The lower molecular mass band present in both the *hap2/HAP2-HA* and *hap2* gametes likely corresponds to the heavy chain of the IgG used for immunoprecipitation.

(B) α-HAP2$_{168-190}$ inhibits gamete fusion when pre-incubated with activated *HAP2-HA minus* gametes, but not WT *plus* gametes. Fusion assays in (B) and (C) were performed at least as duplicates with two counts on fusion rate in each experiment, and results are shown as mean ± SD. Controls had no antibody.

(C) The HAP2$_{168-190}$ peptide (i), but not a control peptide (iv), as well as recombinant HAP2e (ii), neutralized the fusion-inhibiting activity of α-HAP2$_{168-190}$, and a control IgG had no effect on gamete fusion (iii).

(D) HAP2-HA localization at the *minus* mating structure was unaffected by pre-incubation of gametes with α-HAP2$_{168-190}$.

See also Figure S2.

immunolocalization experiments using an anti-HA antibody showed that HAP2-HA on α-HAP2$_{168-190}$-treated *minus* gametes remained at the mating structure (Figure 2D), indicating that α-HAP2$_{168-190}$ did not alter HAP2 localization, but directly interfered with its function. These functional studies with *Chlamydomonas* mutant gametes and the anti-peptide antibody indicated that in its native conformation on live gametes, the 168–190 segment of HAP2 is accessible at the protein surface and its integrity and availability are essential for fusion of gametes.

Recombinant Monomeric HAP2e Trimerizes after Insertion into Lipid Bilayers

We used a *Drosophila* expression system to produce a soluble form of *C. reinhardtii* HAP2 aa 23–592 (comprising almost the entire ectodomain, Figure 1B) and purified it to homogeneity (see STAR Methods; Figure S2). Analysis by size-exclusion chromatography (SEC) and multi-angle static laser light-scattering (MALLS; Figure S2) showed that the protein behaved as a monomer in solution (fraction labeled HAP2e), but had a tendency to

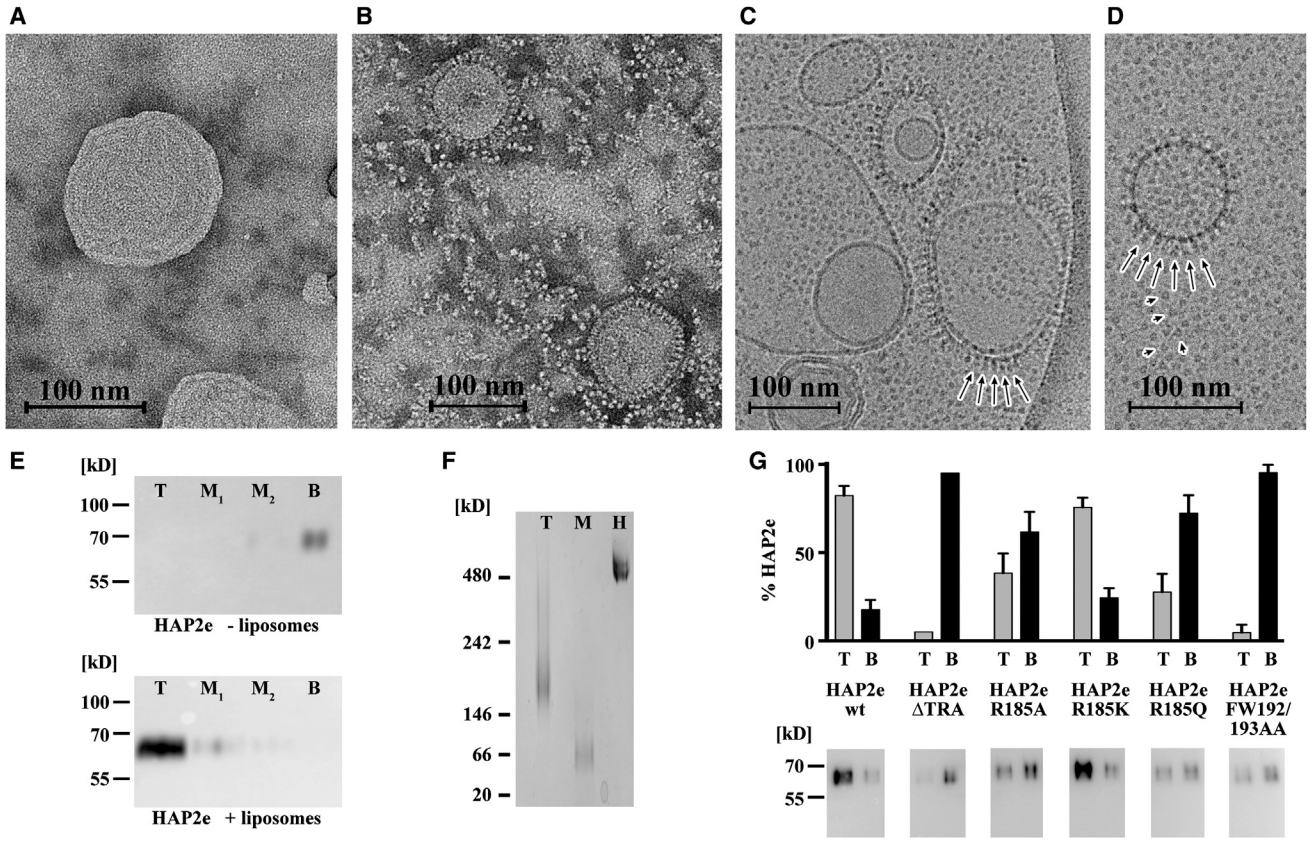


Figure 3. HAP2 Inserts into Liposomes as Trimer

(A–D) Membrane insertion of HAP2e. Electron micrographs of liposomes incubated in the absence (A) and presence (B–D) of *C. reinhardtii* HAP2e and analyzed by negative staining (A and B) or cryo-EM (C and D). Liposomes decorated with HAP2e display protein projections at the surface, contrasting with the smooth liposome surface in the absence of protein (A). The protein projections (some of which are indicated by arrows in B and C) are shaped as rods with a tapered end toward the membrane. The background is coated with proteins not bound to liposomes, which appear to have 3-fold symmetry (arrowheads in D) as they align with the long axis perpendicular to the grid, most likely to expose the hydrophobic end of rod at the water-air interface.

(E) Analysis of HAP2e membrane insertion by co-flotation in sucrose gradients in the absence (top) and presence (bottom) of liposomes, followed by immunoblot detection of HAP2e in the top (T), middle (M1, M2), and bottom (B) fractions.

(F) Native gel analysis of HAP2e solubilized from the top fraction (T), purified HAP2e (M), and hexameric HAP2eh (H; see Figure S2).

(G) HAP2e mutant membrane insertion was analyzed as in E and the amount of HAP2e mutants present in the top (T, gray) and bottom gradient fraction (B, black) is shown as mean ± SD.

See also Figure S3.

oligomerize with time (especially under high ionic strength conditions) to elute at a volume corresponding roughly to hexamers (HAP2eh fraction in Figure S2). The purified protein from the HAP2e monomeric fraction efficiently neutralized the fusion-inhibition potential of the α-HAP2$_{168\text{-}190}$ antibody (Figure 2C), indicating that at least part of the 168–190 segment is exposed in HAP2e and accessible to the antibody.

To detect membrane insertion, we incubated the recombinant protein with liposomes of a standard lipid composition (see STAR Methods) and monitored binding by co-flotation on a sucrose gradient (Figure 3E) followed by immunoblot detection of lipid-inserted HAP2e using a monoclonal antibody raised against HAP2e (mAb K3; see Figure S2; STAR Methods). We observed efficient co-flotation of the monomeric HAP2e fraction (Figure 3E), but not of the multimeric HAP2eh fraction (not shown). Electron microscopy analysis showed that HAP2e decorated the liposome surface as projecting rods about 12 nm long (Figures 3B–3D), which are similar to those formed by viral class II fusion proteins in their trimeric, post-fusion form, such as the alphavirus E1 protein (Gibbons et al., 2003). The size and shape of membrane-bound HAP2e suggested that it had also oligomerized upon membrane insertion. Of note, 3-fold symmetry was apparent in some top views of unbound proteins present in the background (Figure 3D, arrowheads). We confirmed that HAP2e had indeed trimerized upon membrane insertion by detergent-solubilizing it from the liposomes and analyzing it by native PAGE (Figure 3F) and by SEC (Figure S3). These results indicated that HAP2e behaves similarly to the alphaviruses and flaviviruses class II proteins, with membrane insertion concomitant with trimerization of a monomeric pre-fusion form (Klimjack et al., 1994; Stiasny et al., 2002), except that HAP2e did not require an acidic environment for lipid binding and trimerization.

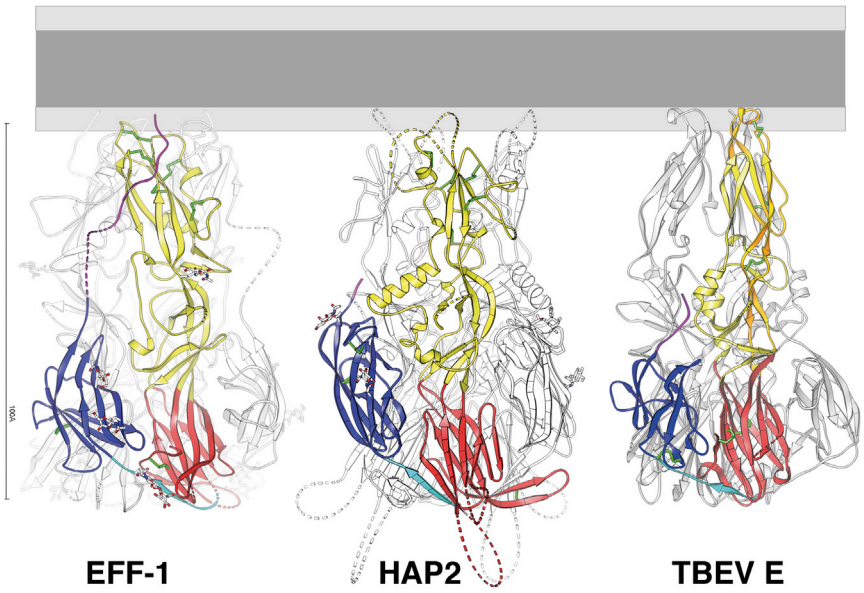


Figure 4. Comparison with Class II Fusion Proteins of Known Structure

Ribbon representation of the *C. reinhardtii* HAP2 trimer shown next to the class II fusion proteins E of the tick-borne encephalitis virus (TBEV) and the *C. elegans* developmental fusogen EFF-1 in their post-fusion trimeric conformation. A "fused" membrane is diagrammed above, roughly to scale, with aliphatic and hydrophilic layers in dark and light gray, respectively. In each case, the trimer subunit in the foreground is colored according to the class II convention: red, yellow, and blue for domains I, II, and III, respectively; magenta and cyan for the stem and the domain I-III linker. Disordered regions are shown as dashed tubes, disulfide bonds as green sticks, and N- and O-linked glycan chains as ball-and-stick colored according to atom type.

See also Figure S5 and Table S1.

This difference is in line with the ability of HAP2 to induce gamete fusion in the extracellular environment, whereas alpha- and flaviviruses require the acidic environment of an endosome for fusion.

Because the co-flotation experiments demanded relatively high amounts of purified recombinant protein for detection, it was impractical to assess flotation inhibition by the undiluted polyclonal α-HAP2$_{168\text{-}190}$ antibody. But we tested the effect of the mutations described in Figure 1 by recombinantly expressing the mutant HAP2 ectodomains. These experiments showed that the mutations that impaired gamete fusion also affected the co-flotation capacity of the mutant HAP2e (Figure 3G). The $R_{185}K$ HAP2e mutant co-floated with liposomes as efficiently as wild-type, whereas those in which the charge at position 185 was removed—the $R_{185}A$ and $R_{185}Q$ mutants—co-floated poorly (Figure 3G). Co-flotation with the liposomes was also essentially abrogated with the HAP2e-ΔTRA and HAP2e-$F_{192}A$-$W_{193}A$ mutants. The combination of the in vivo (Figures 1 and 2) and in vitro (Figure 3) analyses indicates that HAP2 has the capacity to directly interact with membranes and that altering the conserved residues of the HHpred-identified segment affects this interaction.

HAP2 Crystal Structure and Overall Organization of the HAP2 Trimer

We obtained crystals of HAP2e diffracting to a maximum resolution of 3.3 Å. Details of the crystallization and structure determination are described in the STAR Methods, and the crystallographic statistics are listed in Table S1. We determined the X-ray structure by the single isomorphous replacement with anomalous scattering (SIRAS) method using a PtCl4 derivative. The experimental electron density map allowed the tracing of 442 amino acids out of 569 in the HAP2e expression construct (aa 23–592; Figure 1B) from amino acids 24 to 581 with internal breaks at several disordered loops (listed in the STAR Methods).

The atomic model of *C. reinhardtii* HAP2 revealed a trimer with unambiguous structural homology to class II fusion proteins, featuring the three characteristic β sheet-rich domains, termed I, II, and III, arranged in a "hairpin" conformation typical of class II fusion proteins in their post-fusion form (Bressanelli et al., 2004; DuBois et al., 2013; Gibbons et al., 2004; Guardado-Calvo et al., 2016; Halldorsson et al., 2016; Luca et al., 2013; Modis et al., 2004; Willensky et al., 2016) (Figures 4 and S4). The overall shape of the HAP2e trimer is totally compatible with the projections from the liposomes observed in the electron micrographs, with the predicted membrane interacting region at the tapered end of the rods (Figures 3C and 3D). The monomeric HAP2e (Figure S2) used to grow the crystals obviously underwent the same conformational change to form trimers observed during insertion into liposomes (Figure 3). A similar oligomeric rearrangement during crystallization was reported previously for the dengue virus E protein under acidic pH conditions (Klein et al., 2013; Nayak et al., 2009). In the case of HAP2, acidification is not required; the trigger for the rearrangement to the post-fusion form is not known.

The comparison with class II fusion protein trimers of known structure defines the membrane-facing side (top in Figure 4) and the membrane distal end (bottom side) of the HAP2 trimer, in agreement with the shape observed in the electron micrographs (Figures 3B–3D). The Dali server (Holm and Park, 2000) yielded *Z* scores ranging from 9 to 16 for up to ~343 Cα atoms of class II fusion proteins of viral and cellular origin (Table 1), in the same range as those obtained in previous comparisons among known class II fusion proteins (Pérez-Vargas et al., 2014).

HAP2e Protomer

Domain I in HAP2 is a β sandwich of about 200 aa with ten β strands in the two apposed β sheets: the $A_0B_0I_0H_0G_0$ sheet is buried and the $J_0C_0D_0E_0F_0$ sheet is exposed in the trimer, and we refer to them as inner and outer sheets (Figures 5A and 5D). A specific feature of the HAP2 domain I is that the A_0B_0 β-hairpin is long and projects out of the inner sheet to augment

Table 1. Dali Scores Comparing the *C. reinhardtii* HAP2 Ectodomain with Class II Proteins of Known Structure

		Z[a]	N/N^T	σ	I
Alphavirus E1	Semliki forest virus (PDB: 1RER; Gibbons et al., 2004)	13.8	314/391	4.5	10
Flavivirus E	St. Louis encephalitis virus (PDB: 4FG0; Luca et al., 2013)	9.7	291/386	5.6	8
	Tick-borne encephalitis virus (PDB: 1URZ; Bressanelli et al., 2004)	10.0	295/386	5.5	7
	Dengue virus serotype 1 (PDB: 3G7T; Nayak et al., 2009)	9.6	285/380	6.1	8
	Dengue virus serotype 2 (PDB: 1OK8; Modis et al., 2004)	9.7	297/379	5.7	8
C. elegans EFF-1	(PDB: 4OJE; Pérez-Vargas et al., 2014)	13.0	339/457	4.2	7
Rubella Virus E1[b]	(PDB: 4ADI; DuBois et al., 2013)	6.9	247/425	5.1	7
Bunyavirus Gc	Severe fever with thrombocytopenia syndrome virus (PDB: 5G47; Halldorsson et al., 2016)	15.7	343/428	4.3	10

[a]For reference, Z scores below two are meaningless, whereas values of around 50 are obtained when structures of the same protein from two different crystal forms are compared (serving as reference for proteins of roughly the same size). N/N^T, indicates the number of aligned residues (N) compared to the total residues in the alignment (N^T). σ is the root-mean-squared deviation between Cα atoms (in Å). I, indicates % amino acid identity after the alignment.

[b]Comparisons with RV E1 (in which domain III is swapped) were performed with coordinates in which domain III was manually ''unswapped'' to make a trimer organized in the same way as EFF-1 or flavivirus E. Without this operation, the Dali score was based on fewer residues and the *Z* score value poorer.

the outer sheet of the adjacent subunit, where strand A_0 runs parallel to J_0 (Figure S5). These additional inter-subunit main chain β-interactions are likely to confer extra stability to the post-fusion trimer. At the membrane-distal, bottom face of the trimer, the outer β sheet projects very long loops, which are disordered in the crystal (dashed tubes in Figures 4, 5, S4, and S5). The C_0D_0 loop projection displays six additional cysteine residues (presumably forming three disulfide bonds) in an insertion that is present essentially only in algal HAP2 (Liu et al., 2008) (Figure 5D); the E_0F_0 loop projection is also an insertion with four potential N-linked glycosylation sites.

Domain III is made of about 130 aa and has an immunoglobulin-like fold, with seven β strands in two sheets (labeled ABE and GFCC'; Figures 5A, 5D, and S4). Although the β sheets are longer than in other class II fusion proteins (Figure 4), giving domain III a more elongated bean shape, it is in a similar location at the side of the trimer (Figure 4) as in other class II fusion proteins of known structure. Domain III packs against two adjacent protomers of the central trimer ''core,'' which is composed of domains I and II arranged parallel to each other and interacting along their length about the 3-fold molecular axis. The latter central trimer interaction is postulated to exist during formation of an ''extended trimeric intermediate'' state of the fusogenic conformational transition (Liao et al., 2010), in which the fusion loops are inserted in the target membrane (*plus* gamete here), while the C-terminal TM segment is anchored in the *minus* gamete mating structure, at the opposite end. The final collapse into the post-fusion hairpin conformation of each protomer in the trimer brings domain III to the sides of this core, projecting the downstream stem and TM regions toward the fusion loop (Liao and Kielian, 2005), in a fusogenic rearrangement of the protein analogous to an umbrella folding inside out. In this final, post-fusion location, domain III buries an area of about 2,100 Å^2 of its surface, divided roughly equally in contacts with the same and with the adjacent subunit (intra- and inter-subunit contacts). The observed contacts therefore can only form after the assembly of the central trimer core, in line with the proposed clamping role of domain III, and resulting in irreversible trimerisation, as proposed for other class II fusion proteins (Liao and Kielian, 2005; Pérez-Vargas et al., 2014).

HAP2 Has Two Putative Fusion Loops between β Strands *c* and *d*

Domain II is the largest domain (roughly 250 aa in total) and is made by two distinct segments emanating from domain I: the D_0E_0 and H_0I_0 strand connections of the outer and inner sheets of domain I, respectively (Figure S4). As in all class II proteins, the domain I proximal region of domain II has a central β sheet (*aefg*; Figures 5 and S4) flanked by additional short helices. The distal ''tip'' of domain II contains β sheet *bdc*, with the β strands running parallel to the molecular 3-fold axis at the distal end of the D_0E_0 segment. The connection between β strands *c* and *d* at the tip of domain II (the *cd* loop) was shown to be the fusion loop in the viral proteins (reviewed in Kielian and Rey [2006]). The *bdc* sheet normally packs against the *ij* β-hairpin (the distal end of the H_0I_0 segment), which in HAP2 maps to the conserved HAP2/GCS1 PFAM domain PF10699) (Figure 5A). Although in HAP2 β strands *i* and *j* are absent, we still refer to this region as the ''*ij* loop'' (Figures 5 and S4).

The *cd* loop in HAP2 is 40 aa long and has an intervening short α-helix in the middle (α0), in contrast to the standard class II fusion proteins from the arthropod-borne viruses such as flaviviruses and alphaviruses (10–15 residues long). In this respect, HAP2 resembles the rubella virus class II fusion protein E1, which has 50 aa in between stands *c* and *d*, with a couple of short intervening α helices as well as an additional β strand which results in two separate fusion loops (DuBois et al., 2013). The presence of the α0 helix in HAP2 also results in two loops (loops 1 and 2 in Figures 5B and 5C), which project outward. Although disordered in the crystal, these two loops are in position to project non-polar residues into the target membrane. The HHpred alignment was indeed quite close and pointed correctly to the *cd* loop (Figure S1). The $HAP2_{168\text{-}190}$ peptide used for immunization spans loop 1 all the way to the end of helix α0, and the fact that the

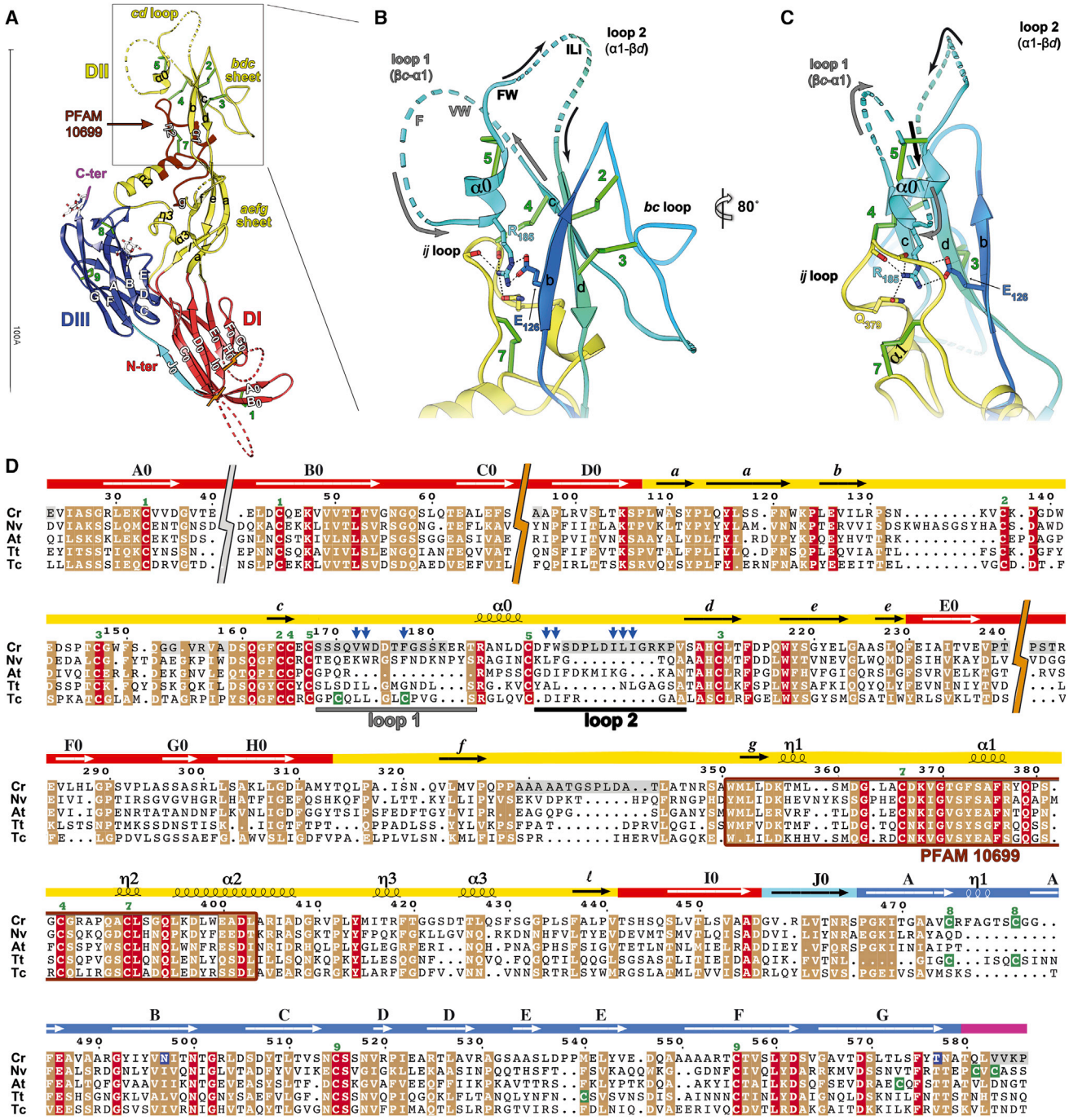


Figure 5. 3D Structure of HAP2

(A) Ribbon diagram of one subunit of the *C. reinhardtii* HAP2 trimer colored as in Figure 4 with the HAP2/GCS1 PFAM domain highlighted in brown and labeled (arrow). Green arabic numbers indicate disulfides and match those in (D).

(B and C) Close-up view of the tip region, highlighting the bipartite *cd*-loop displaying two hydrophobic patches (blue vertical arrows in D) and the structural role of the conserved residues R185 and E126. The HAP2 subunit is ramp-colored from N to C terminus (as in Figure S4) to better distinguish the two distal ends of the polypeptide segments emanating from domain I (D_0E_0 and H_0I_0 connections), which project out the *cd* loop (blue) and the *ij* loop (yellow) regions, respectively. Curved arrows along the two fusion loops show the directionality of the polypeptide chain along loops 1 (gray arrows) and 2 (black arrows). Disulfide 4, together with a network of hydrogen bonds (black dotted lines) involving the side chains of R_{185}, E_{126}, Q_{379} and backbone carbonyl groups (B) tightly interconnect the *ij*-loop to the *bdc*-sheet and the central segment of the *cd* loop, marked by helix α0 (C).

(D) Amino acid alignment of the HAP2 ectodomains from *Chlamydomonas reinhardtii* (Cr), the sea anemone *Nematostella vectensis* (Nv), the flowering plant *Arabidopsis thaliana* (At), the ciliated protozoan *Tetrahymena thermophila* (Tt), and the protist *Trypanosoma cruzi* (Tc) drawn with ESPript (Gouet et al., 1999) and

(legend continued on next page)

resulting polyclonal antibody blocks fusion is in line with this region being exposed at this end of the molecule. The HAP2 crystal structure is therefore compatible with the mutagenesis data in this region, and with the effect of the α-HAP2$_{168\text{-}190}$ antibody. Moreover, a peptide derived from the corresponding region of *Tetrahymena thermophila* HAP2 was found to display properties typical of a fusion loop (see related paper in *Current Biology*, Pinello et al., 2017), in line with our findings with *C. reinhardtii*.

Arginine 185 Organizes the Target-Membrane-Interacting Region of HAP2

The structure shows that the functionally important residue R_{185} is at the N terminus of the α0 helix and its side chain points away from the two exposed loops and toward the core of the HAP2 trimer. R_{185} makes a salt bridge and bidentate hydrogen bonds with the side chain of the strictly conserved E_{126} in β strand *b* (Figures 5B and 5C). Furthermore, the R_{185} side chain is at the core of a network of interactions stabilizing the main-chain conformation of the *ij* loop. This network involves main chain atoms of F_{376}, G_{382} and R_{385}, together with the strictly conserved Q_{379} side chain, which are part of the PF10699 signature segment that allowed the identification of HAP2 in widely disparate organisms. The structure now shows that the conserved pattern of disulfide bonds of this signature element is required for the *ij* loop to adopt a convoluted fold acting as a framework underpinning the cd loop by interaction with the conserved R_{185} so that the two fusion loops project out at the membrane-interacting region. The latter region, in contrast, is variable and has multiple deletions and insertions in the various orthologs, some insertions being quite long (Figure 1C). It is possible that the differences in membrane-interacting regions of HAP2s across the broad spectrum of eukaryotic organisms reflect evolutionary adaptations required for fusion with different target gametes. We note that the fusion loop in viral class II proteins is in general the most conserved segment of the protein within orthologs from a given virus genus, most likely because the same residues are also required for inter-subunit interactions in their pre-fusion form on infectious particles. But this comparison is not necessarily informative, since the analyzed HAP2 proteins span eukaryotic taxa that are much more distantly related than are viruses within a given genus. The pre-fusion conformation of HAP2 remains unknown, however, and it is possible that the non-polar residues of the two fusion loops are maintained unexposed until the time of fusion. The conserved interaction of the R_{185} side chain with that of the E_{126} residue in the tip domain may allow exposure of the two fusion loops only after a conformational change. In this context, the fusion-inhibiting α-HAP2$_{168\text{-}190}$ antibody would bind only after a conformational change that exposes the fusion loop, but this remains to be explored.

Virus-Cell Gene Exchanges during the Evolution of Eukaryotes

The identical topological arrangement of secondary, tertiary, and quaternary structure elements of HAP2 with the viral class II proteins can only be explained by postulating a common ancestor. Indeed, the probability of convergence to the observed complex fold from independent origins, to result in two proteins displaying exactly the same topological arrangement throughout the entire ectodomain (as shown in Figure S4) is extremely low and can be considered negligible. Nature is parsimonious, and once a protein required for a complex function such as membrane fusion becomes available, the corresponding gene is used over and over again, most likely transferred via horizontal gene exchanges. This concept is supported by the observation that only three structural classes of viral fusion proteins have been observed so far, in spite of the enormous variety of known viruses. As one example, the membrane fusion proteins of the herpesviruses, rhabdoviruses and baculoviruses, which are otherwise totally unrelated viruses, were shown to be homologous by structural studies (Heldwein et al., 2006; Kadlec et al., 2008; Roche et al., 2006). These fusion proteins most certainly derived from a distant common ancestor (class III proteins) (Backovic and Jardetzky, 2011), whose genes must have been acquired via horizontal exchanges. Within eukaryotic organisms, only a few other cell-cell fusion proteins have been positively identified. Although the myoblast myomaker, a seven-pass transmembrane protein that governs fusion of myoblasts to form myotubes (Millay et al., 2013, 2016), has no obvious relation to viruses, the proteins involved in the two other cell-cell fusion events that have been characterized at the molecular level are virus related. Cytotrophoblast fusion in mammals during placenta formation (Blaise et al., 2003; Holm and Park, 2000) and epidermal cell-cell fusion in nematodes to form syncytia (Mohler et al., 2002) both use fusion proteins also found in viruses. In the first case, the class I fusion protein involved is clearly derived from an endogenous retroviral element (Denner, 2016). It is also possible that a similar process involving retrotranscription may have taken place in the case of the class II *Caenorhabditis elegans* fusion protein EFF-1 (Pérez-Vargas et al., 2014), as retroviruses of nematodes have been found to have an envelope protein related to that of the phleboviruses (Frame et al., 2001; Malik et al., 2000), which have class II fusion proteins (Dessau and Modis, 2013). These observations suggest that retro-transcription, followed by integration into the genome, may have been an important pathway for gene exchanges between viruses and eukaryotic cells.

Concluding Remarks

HAP2 is present in organisms responsible for several of the globe's most devastating human diseases, including *Plasmodium*, *Trypanosoma*, and *Toxoplasma*. Many arthropods that are vectors of human diseases or are agricultural pests, such

color-coded as in Figure 1C, but with a gray background denoting segments disordered in the structure. Indicated above the alignment are the secondary structure elements. For clarity the breaks at the two long disordered loops C_0D_0 and E_0F_0 in *C. reinhardtii* HAP2 domain I are indicated by an orange flash and an insertion in *N. vectensis* HAP2 in the A_0B_0 connection by a gray flash. The sites of attached N- and O-linked glycans are indicated in blue. The PFAM 10699 segment is boxed. Vertical blue arrows above the sequence point to non-polar residues in loop 1 and loop 2 that can potentially insert into the outer leaflet of the target membrane.
See also Figure S4.

as insects and ticks, also possess HAP2. A strategy that used recombinant *Plasmodium* HAP2 fragments to induce transmission-blocking immunity in mice has been reported previously (Blagborough and Sinden, 2009; Miura et al., 2013), but the expression systems were not sufficient for viable clinical developments. Our results suggest that the use of a peptide spanning the HAP2 fusion loops as immunogen might be sufficient to induce transmission-blocking immunity, similar to the antibodies we obtained here against *Chlamydomonas* HAP2.

In conclusion, our data now open the way to a full mechanistic characterization of gamete fusion induced by HAP2 and raise new questions, including the identification of the trigger for the HAP2 fusogenic conformational change, the structure of the pre-fusion form(s) of HAP2, and its organization on the mating structure of the *minus* gamete membrane. Evolution through hundreds of millions of years may have led the different taxa to develop alternative solutions, and the metastable pre-fusion form of HAP2 may be organized differently in the multiple organisms in which it is present, as shown for viruses of different families. In contrast, the post-fusion conformation described here appears as a universal feature of class II fusion proteins.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - *Chlamydomonas reinhardtii* Cells and Cell Culture
- METHOD DETAILS
  - Fusion Inhibition Assay Using an anti $HAP2_{168-190}$ Peptide Antibody
  - Immunolocalization, Immunoprecipitation, and Trypsin Experiments
  - Generating Chlamydomonas Cells Expressing Mutant Forms HAP2-HA
  - Expression and Purification of the Soluble HAP2 Ectodomain
- SEC-MALLS
  - Production of Monoclonal Antibodies Targeting *C. reinhardtii* HAP2
  - Liposome Co-flotation Experiments
  - Electron Microscopy
  - Native Gel Analysis
  - Crystallization and Structure Determination
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2017.01.024.

AUTHOR CONTRIBUTIONS

W.J.S. and Y.L. conceived the in vivo experiments with full-length HAP2, and T.K. and F.A.R. conceived the in vitro experiments with the recombinant HAP2 ectodomain. J.P. and N.G. performed the HAP2 homology search and generated alignments. W.J.S., Y.L., J.P., and N.G. designed the mutant HAP2 proteins and selected the peptide used to produce the fusion-blocking antibody. W.J.S., Y.L., and W.L. designed the experiments with *Chlamydomonas* gametes; Y.L. and W.L. performed the gamete fusion assays with HAP2 mutants and antibodies and interpreted them with W.J.S. J.F. and A.M. made the constructs, expressed and crystallized the HAP2e protein, and J.F. optimized the crystals for diffraction. F.T. raised and isolated monoclonal antibodies targeting *C. reinhardtii* HAP2. J.F. and T.K. collected data and determined the HAP2 X-ray structure and G.B. helped devise a collection strategy for derivative X-ray data. J.F. and M.A.T. demonstrated insertion of HAP2 into liposomes, and G.P.A. performed electron microscopy on liposomes. T.K., F.A.R., and W.J.S. supervised experimental work and wrote the manuscript with input from all other authors.

ACKNOWLEDGMENTS

We thank Francis-André Wollmann and Sandrine Bujaldon for discussions and help with initial experiments; Fredrick Grinnell, Saikat Mukhopadhyay, Michael Henne, and Margaret Phillips of UT Southwestern for helpful discussions; the crystallization platform of the Pasteur Proteopole for technical help; Remi Fronzes for help with native PAGE experiments; Vincent Olieric and the staff of synchrotron beamlines PX-III at the Swiss Light Source, Proxima-1 and -2 at SOLEIL and ID29 and ID30-3 at the European Synchrotron Radiation Facility for help during data collection; members of the Rey and Snell labs for discussions; and M. Nilges and the Equipement d'excellence CACSICE for providing the Falcon II direct electron detector. F.A.R. acknowledges funding from the ERC Advanced grant project (340371) CelCelFus for this work, which also used general support from Institut Pasteur and CNRS. W.J.S. was supported by a grant from the NIH (GM56778) and acknowledges funding from the UTSW HI/HR Program. N.V.G. is funded, in part, by a grant from the NIH (GM094575) and the Welch Foundation (I-1505). J.F. benefitted from an Allocation ministérielle pour l'Ecole Polytechnique AMX.

Received: November 17, 2016
Revised: January 3, 2017
Accepted: January 19, 2017
Published: February 23, 2017

REFERENCES

Backovic, M., and Jardetzky, T.S. (2011). Class III viral membrane fusion proteins. Adv. Exp. Med. Biol. *714*, 91–101.

Belzile, O., Hernandez-Lara, C.I., Wang, Q., and Snell, W.J. (2013). Regulated membrane protein entry into flagella is facilitated by cytoplasmic microtubules and does not require IFT. Curr. Biol. *23*, 1460–1465.

Bianchi, E., Doe, B., Goulding, D., and Wright, G.J. (2014). Juno is the egg Izumo receptor and is essential for mammalian fertilization. Nature *508*, 483–487.

Blagborough, A.M., and Sinden, R.E. (2009). Plasmodium berghei HAP2 induces strong malaria transmission-blocking immunity in vivo and in vitro. Vaccine *27*, 5187–5194.

Blaise, S., de Parseval, N., Bénit, L., and Heidmann, T. (2003). Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. Proc. Natl. Acad. Sci. USA *100*, 13013–13018.

Bressanelli, S., Stiasny, K., Allison, S.L., Stura, E.A., Duquerroy, S., Lescar, J., Heinz, F.X., and Rey, F.A. (2004). Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. EMBO J. *23*, 728–738.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, P., Roversi, P., Sharff, A., Smart, O., Vonrhein, C., et al. (2010). BUSTER, version 2.9 (Global Phasing Ltd.).

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010).

MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Cole, E.S., Cassidy-Hanley, D., Fricke Pinello, J., Zeng, H., Hsueh, M., Kolbin, D., Ozzello, C., Giddings, T., Jr., Winey, M., and Clark, T.G. (2014). Function of the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate. Curr. Biol. *24*, 2168–2173.

Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta. Crystallogr. D. Biol. Crystallogr. *50*, 760–763.

Denner, J. (2016). Expression and function of endogenous retroviruses in the placenta. APMIS *124*, 31–43.

Dessau, M., and Modis, Y. (2013). Crystal structure of glycoprotein C from Rift Valley fever virus. Proc. Natl. Acad. Sci. USA *110*, 1696–1701.

Dong, A., Xu, X., Edwards, A.M., Chang, C., Chruszcz, M., Cuff, M., Cymborowski, M., Di Leo, R., Egorova, O., Evdokimova, E., et al.; Midwest Center for Structural Genomics; Structural Genomics Consortium (2007). In situ proteolysis for protein crystallization and structure determination. Nat. Methods *4*, 1019–1021.

DuBois, R.M., Vaney, M.-C., Tortorici, M.A., Kurdi, R.A., Barba-Spaeth, G., Krey, T., and Rey, F.A. (2013). Functional and evolutionary insight from the crystal structure of rubella virus protein E1. Nature *493*, 552–556.

Ebchuqin, E., Yokota, N., Yamada, L., Yasuoka, Y., Akasaka, M., Arakawa, M., Deguchi, R., Mori, T., and Sawada, H. (2014). Evidence for participation of GCS1 in fertilization of the starlet sea anemone Nematostella vectensis: implication of a common mechanism of sperm-egg fusion in plants and animals. Biochem. Biophys. Res. Commun. *451*, 522–528.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. *62*, 72–82.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., et al. (2016). The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. *44*(D1), D279–D285.

Frame, I.G., Cutfield, J.F., and Poulter, R.T. (2001). New BEL-like LTR-retrotransposons in Fugu rubripes, Caenorhabditis elegans, and Drosophila melanogaster. Gene *263*, 219–230.

Gibbons, D.L., Erk, I., Reilly, B., Navaza, J., Kielian, M., Rey, F.A., and Lepault, J. (2003). Visualization of the target-membrane-inserted fusion protein of Semliki Forest virus by combined electron microscopy and crystallography. Cell *114*, 573–583.

Gibbons, D.L., Vaney, M.C., Roussel, A., Vigouroux, A., Reilly, B., Lepault, J., Kielian, M., and Rey, F.A. (2004). Conformational change and protein-protein interactions of the fusion protein of Semliki Forest virus. Nature *427*, 320–325.

Gouet, P., Courcelle, E., Stuart, D.I., and Métoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics *15*, 305–308.

Guardado-Calvo, P., Bignon, E.A., Stettner, E., Jeffers, S.A., Pérez-Vargas, J., Pehau-Arnaudet, G., Tortorici, M.A., Jestin, J.L., England, P., Tischler, N.D., and Rey, F.A. (2016). Mechanistic insight into bunyavirus-induced membrane fusion from structure-function analyses of the hantavirus envelope glycoprotein Gc. PLoS Pathog. *12*, e1005813.

Halldorsson, S., Behrens, A.J., Harlos, K., Huiskonen, J.T., Elliott, R.M., Crispin, M., Brennan, B., and Bowden, T.A. (2016). Structure of a phleboviral envelope glycoprotein reveals a consolidated model of membrane fusion. Proc. Natl. Acad. Sci. USA *113*, 7154–7159.

Harrison, S.C. (2015). Viral membrane fusion. Virology *479-480*, 498–507.

Heldwein, E.E., Lou, H., Bender, F.C., Cohen, G.H., Eisenberg, R.J., and Harrison, S.C. (2006). Crystal structure of glycoprotein B from herpes simplex virus 1. Science *313*, 217–220.

Holm, L., and Park, J. (2000). DaliLite workbench for protein structure comparison. Bioinformatics *16*, 566–567.

Johnson, M.A., von Besser, K., Zhou, Q., Smith, E., Aux, G., Patton, D., Levin, J.Z., and Preuss, D. (2004). Arabidopsis hapless mutations define essential gametophytic functions. Genetics *168*, 971–982.

Kabsch, W. (1988). Automatic indexing of rotation diffraction patterns. J. Appl. Cryst. *21*, 67–72.

Kadlec, J., Loureiro, S., Abrescia, N.G., Stuart, D.I., and Jones, I.M. (2008). The postfusion structure of baculovirus gp64 supports a unified view of viral fusion machines. Nat. Struct. Mol. Biol. *15*, 1024–1030.

Kawai-Toyooka, H., Mori, T., Hamaji, T., Suzuki, M., Olson, B.J.S.C., Uemura, T., Ueda, T., Nakano, A., Toyoda, A., Fujiyama, A., and Nozaki, H. (2014). Sex-specific posttranslational regulation of the gamete fusogen GCS1 in the isogamous volvocine alga *Gonium pectorale*. Eukaryot. Cell *13*, 648–656.

Kielian, M., and Rey, F.A. (2006). Virus membrane-fusion proteins: more than one way to make a hairpin. Nat. Rev. Microbiol. *4*, 67–76.

Klein, D.E., Choi, J.L., and Harrison, S.C. (2013). Structure of a dengue virus envelope protein late-stage fusion intermediate. J. Virol. *87*, 2287–2293.

Klimjack, M.R., Jeffrey, S., and Kielian, M. (1994). Membrane and protein interactions of a soluble form of the Semliki Forest virus fusion protein. J. Virol. *68*, 6940–6946.

Krey, T., d'Alayer, J., Kikuti, C.M., Saulnier, A., Damier-Piolle, L., Petitpas, I., Johansson, D.X., Tawar, R.G., Baron, B., Robert, B., et al. (2010). The disulfide bonds in glycoprotein E2 of hepatitis C virus reveal the tertiary organization of the molecule. PLoS Pathog. *6*, e1000762.

Liao, M., and Kielian, M. (2005). Domain III from class II fusion proteins functions as a dominant-negative inhibitor of virus membrane fusion. J. Cell Biol. *171*, 111–120.

Liao, M., Sánchez-San Martín, C., Zheng, A., and Kielian, M. (2010). In vitro reconstitution reveals key intermediate states of trimer formation by the dengue virus membrane fusion protein. J. Virol. *84*, 5730–5740.

Lillie, F.R. (1913). The Mechanism of Fertilization. Science *38*, 524–528.

Liu, Y., Tewari, R., Ning, J., Blagborough, A.M., Garbom, S., Pei, J., Grishin, N.V., Steele, R.E., Sinden, R.E., Snell, W.J., and Billker, O. (2008). The conserved plant sterility gene HAP2 functions after attachment of fusogenic membranes in *Chlamydomonas* and *Plasmodium* gametes. Genes Dev. *22*, 1051–1068.

Liu, Y., Misamore, M.J., and Snell, W.J. (2010). Membrane fusion triggers rapid degradation of two gamete-specific, fusion-essential proteins in a membrane block to polygamy in *Chlamydomonas*. Development *137*, 1473–1481.

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.

Luca, V.C., Nelson, C.A., and Fremont, D.H. (2013). Structure of the St. Louis encephalitis virus postfusion envelope trimer. J. Virol. *87*, 818–828.

Malik, H.S., Henikoff, S., and Eickbush, T.H. (2000). Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. Genome Res. *10*, 1307–1318.

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

Millay, D.P., O'Rourke, J.R., Sutherland, L.B., Bezprozvannaya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2013). Myomaker is a membrane activator of myoblast fusion and muscle formation. Nature *499*, 301–305.

Millay, D.P., Gamage, D.G., Quinn, M.E., Min, Y.L., Mitani, Y., Bassel-Duby, R., and Olson, E.N. (2016). Structure-function analysis of myomaker domains required for myoblast fusion. Proc. Natl. Acad. Sci. USA *113*, 2116–2121.

Misamore, M.J., Gupta, S., and Snell, W.J. (2003). The *Chlamydomonas* Fus1 protein is present on the mating type plus fusion organelle and required for a critical membrane adhesion event during fusion with minus gametes. Mol. Biol. Cell *14*, 2530–2542.

Miura, K., Takashima, E., Deng, B., Tullo, G., Diouf, A., Moretz, S.E., Nikolaeva, D., Diakite, M., Fairhurst, R.M., Fay, M.P., et al. (2013). Functional comparison

of *Plasmodium falciparum* transmission-blocking vaccine candidates by the standard membrane-feeding assay. Infect. Immun. *81*, 4377–4382.

Modis, Y., Ogata, S., Clements, D., and Harrison, S.C. (2004). Structure of the dengue virus envelope protein after membrane fusion. Nature *427*, 313–319.

Mohler, W.A., Shemer, G., del Campo, J.J., Valansi, C., Opoku-Serebuoh, E., Scranton, V., Assaf, N., White, J.G., and Podbilewicz, B. (2002). The type I membrane protein EFF-1 is essential for developmental cell fusion. Dev. Cell *2*, 355–362.

Mori, T., Kuroiwa, H., Higashiyama, T., and Kuroiwa, T. (2006). Generative cell specific 1 is essential for angiosperm fertilization. Nat. Cell Biol. *8*, 64–71.

Mori, T., Igawa, T., Tamiya, G., Miyagishima, S.Y., and Berger, F. (2014). Gamete attachment requires GEX2 for successful fertilization in *Arabidopsis*. Curr. Biol. *24*, 170–175.

Nayak, V., Dessau, M., Kucera, K., Anthony, K., Ledizet, M., and Modis, Y. (2009). Crystal structure of dengue virus type 1 envelope protein in the postfusion conformation and its implications for membrane fusion. J. Virol. *83*, 4338–4344.

Ning, J., Otto, T.D., Pfander, C., Schwach, F., Brochet, M., Bushell, E., Goulding, D., Sanders, M., Lefebvre, P.A., Pei, J., et al. (2013). Comparative genomics in *Chlamydomonas* and *Plasmodium* identifies an ancient nuclear envelope protein family essential for sexual reproduction in protists, fungi, plants, and vertebrates. Genes Dev. *27*, 1198–1215.

Okabe, M. (2013). The cell biology of mammalian fertilization. Development *140*, 4471–4479.

Pérez-Vargas, J., Krey, T., Valansi, C., Avinoam, O., Haouz, A., Jamin, M., Raveh-Barak, H., Podbilewicz, B., and Rey, F.A. (2014). Structural basis of eukaryotic cell-cell fusion. Cell *157*, 407–419.

Pinello, J.F., Lai, A.L., Millet, J.K., Cassidy-Hanley, D., Freed, J.H., and Clark, T.G. (2017). Structure-function studies link class II viral fusogens and the ancestral gamete fusion protein HAP2. Curr. Biol. *27*. Published online February 23, 2017. http://dx.doi.org/10.1016/j.cub.2017.01.049.

Roche, S., Bressanelli, S., Rey, F.A., and Gaudin, Y. (2006). Crystal structure of the low-pH form of the vesicular stomatitis virus glycoprotein G. Science *313*, 187–191.

Sinden, R.E., Canning, E.U., and Spain, B. (1976). Gametogenesis and fertilization in *Plasmodium yoelii nigeriensis*: a transmission electron microscope study. Proc. R. Soc. Lond. B Biol. Sci. *193*, 55–76.

Snell, W.J., and Goodenough, U.W. (2009). Flagellar adhesion, flagellar-generated signaling, and gamete fusion during mating. In The *Chlamydomonas* Sourcebook, E.H. Harris, D.B. Stern, and G.B. Witman, eds. (Elsevier), pp. 369–394.

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.

Sprunck, S., Rademacher, S., Vogler, F., Gheyselinck, J., Grossniklaus, U., and Dresselhaus, T. (2012). Egg cell-secreted EC1 triggers sperm cell activation during double fertilization. Science *338*, 1093–1097.

Steele, R.E., and Dana, C.E. (2009). Evolutionary history of the HAP2/GCS1 gene and sexual reproduction in metazoans. PLoS ONE *4*, e7680.

Stiasny, K., Allison, S.L., Schalich, J., and Heinz, F.X. (2002). Membrane interactions of the tick-borne encephalitis virus fusion protein E at low pH. J. Virol. *76*, 3784–3790.

van Dijk, M.R., Janse, C.J., Thompson, J., Waters, A.P., Braks, J.A., Dodemont, H.J., Stunnenberg, H.G., van Gemert, G.J., Sauerwein, R.W., and Eling, W. (2001). A central role for P48/45 in malaria parasite male gamete fertility. Cell *104*, 153–164.

von Besser, K., Frank, A.C., Johnson, M.A., and Preuss, D. (2006). *Arabidopsis* HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. Development *133*, 4761–4769.

Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. Methods Mol. Biol. *364*, 215–230.

Weinert, T., Olieric, V., Waltersperger, S., Panepucci, E., Chen, L., Zhang, H., Zhou, D., Rose, J., Ebihara, A., Kuramitsu, S., et al. (2014). Fast native-SAD phasing for routine macromolecular structure determination. Nat. Methods *12*, 131–133.

Willensky, S., Bar-Rogovsky, H., Bignon, E.A., Tischler, N.D., Modis, Y., and Dessau, M. (2016). Crystal structure of glycoprotein C from a hantavirus in the post-fusion conformation. PLoS Pathog. *12*, e1005948.

Wong, J.L., and Johnson, M.A. (2010). Is HAP2-GCS1 an ancestral gamete fusogen? Trends Cell Biol. *20*, 134–141.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Polyclonal rabbit α-HAP2168-190 peptide Ab	This paper	N/A
Monoclonal mouse α-HAP2e Ab K3	This paper	N/A
Chemicals, Peptides, and Recombinant Proteins		
$HAP2_{168-190}$ (SSSQVWDDTFGSSKERTRANLDC)	YenZym Antibodies	N/A
$HAP2_{control}$ (CTQPPRPPWPPRPPPAPPPS)	YenZym Antibodies	N/A
Common lab reagents	N/A	N/A
Deposited Data		
C. reinhardtii HAP2e structure	This paper	PDB: 5MF1
Experimental Models: Cell Lines		
D. melanogaster: Cell line S2: S2	Thermo-Fisher	Cat # R690-07
Experimental Models: Organisms/Strains		
Chlamydomonas reinhardtii: strain 21 gr: mt+	Chlamydomonas Culture Collection	CC-1690
Chlamydomonas reinhardtii: strain *40D4*: hap2 mating type *minus* mutant	Chlamydomonas Culture Collection	CC-4552
Chlamydomonas reinhardtii: strain *HAP2-HA*: *HAP2-HA* plasmid-rescued *40D4* strain	Chlamydomonas Culture Collection	CC-5295
Recombinant DNA		
pMT/BiP/TwinStrep	Krey et al., 2010	N/A
pMT/BiP/HAP2e23-592TwinStrep	This paper	N/A
Software and Algorithms		
XDS	Kabsch, 1988	http://xds.mpimf-heidelberg.mpg.de/
Pointless	Evans, 2006	http://www.ccp4.ac.uk/
CCP4	Collaborative Computational Project, 1994	http://www.ccp4.ac.uk/
autoSHARP	Vonrhein et al., 2007	https://www.globalphasing.com
Phaser	McCoy et al., 2007	http://www.phaser.cimr.cam.ac.uk/index.php/Phaser_Crystallographic_Software
Coot	Emsley et al., 2010	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
AutoBuster	Bricogne et al., 2010	https://www.globalphasing.com
MolProbity	Chen et al., 2010	http://molprobity.biochem.duke.edu/
HHPred	Söding et al., 2005	https://toolkit.tuebingen.mpg.de/hhpred
DALI	Holm and Park, 2000	http://ekhidna.biocenter.helsinki.fi/dali_server/start

CONTACT FOR REAGENT AND RESOURCE SHARING

Requests for resources and reagents should be directed to Félix A. Rey (felix.rey@pasteur.fr).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Chlamydomonas reinhardtii Cells and Cell Culture

Wild-type *Chlamydomonas* strains *21 gr* (mating type *plus*; mt+; CC-1690), *40D4* (*hap2* mating type *minus* mutant (Liu et al., 2015)) and *HAP2-HA* (*HAP2-HA* plasmid-rescued *hap2* strain (Liu et al., 2015)) were used and are available from the *Chlamydomonas* Culture Collection. Vegetative growth of cells and induction of gametogenesis in gamete medium (M-N) were as described before (Liu, et al., 2008). Gametes were activated with dibutyryl-cAMP (db-cAMP) by incubation with 15 mM db-cAMP and 0.15 mM papav-

erine for 0.5 hr in N-free medium (Ning et al., 2013). Gamete fusion was assessed by determining the number of cells that had formed zygotes after being mixed with wild-type *plus* gametes for 30 min and was expressed as percent fusion using the following equation: (2 × number of zygotes)/[(2 × number of zygotes)+(number of unfused gametes)] × 100 (Liu et al., 2015).

METHOD DETAILS

Fusion Inhibition Assay Using an anti HAP2$_{168\text{-}190}$ Peptide Antibody

A polyclonal antibody against HAP2 peptide SSSQVWDDTFGSSKERTRANLDC (aa 168-190) was made in rabbits by Yenzym Antibodies (San Francisco, CA; α-HAP2$_{168\text{-}190}$), under oversight by their Institutional Review Board. The antibody was purified on a peptide-conjugated affinity column prepared by the company. To assay for antibody inhibition of gamete fusion, gametes that had been activated with dibutyryl-cAMP for 30 min (Misamore et al., 2003) were washed once with gamete medium and the activated gametes (2 × 10^7 cell/ml) were incubated with 110 ug/ml peptide antibody (pre-dialyzed in gamete medium) for 2 hr followed by mixing with gametes of the opposite mating type gametes for the times indicated in the figure legends. The number of zygotes (detected as cells with 4 rather than 2 cilia) was determined by phase-contrast microscopy. At least 100 cells were counted each time with at least two counts per sample in at least two separate experiments. A rabbit IgG (Sigma) was used as a control antibody.

To test the capacity of peptides or HAP2e to neutralize the fusion-inhibiting properties of α-HAP2$_{168\text{-}190}$, the antibody (220 ug/ml) that had been dialyzed in gamete medium was incubated overnight with control peptide CTQPPRPPWPPRPPPAPPPS (200 ug/ml) (this peptide is encoded by Cre03 g176961 whose transcripts are specific to *minus* gametes and upregulated during gamete activations (Ning et al., 2013), HAP2$_{168\text{-}190}$ peptide (200 ug/ml), or HAP2e protein (200 ug/ml) overnight and then added to activated gametes as above to determine percent gamete fusion.

Immunolocalization, Immunoprecipitation, and Trypsin Experiments

Immunofluorescent staining of gametes was performed as described previously (Belzile et al., 2013). The gametes were fixed in ice cold methanol; the primary antibody was rat anti-HA (Roche) and the secondary was Alexa Fluor 488 Goat Anti-Rat (Invitrogen). Images were captured in DIC or FITC channels as described previously (Liu et al., 2015). For HAP2-HA localization after pre-incubation in α-HAP2$_{168\text{-}190}$ (see below), db-cAMP-activated gametes were incubated with the antibody for 2 hr, washed three times with gamete medium, and stained with HA antibody as above.

For immunoprecipitation, gametes were disrupted by incubation at 4°C for 30 min in RIPA buffer (20 mM Tris, 150 mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% SDS) containing a proteinase inhibitor cocktail (Roche). The samples were centrifuged at 12000 rpm for 30 min, the supernatants were subjected to immunoprecipitation with α-HAP2$_{168\text{-}190}$ and protein A agarose beads using methods described previously (Liu et al., 2010) and the immunoprecipitated proteins were subjected to SDS-PAGE and immunoblotting (Liu et al., 2015). For trypsin treatment 5x10^7 live gametes/ml were incubated in 0.05% freshly prepared trypsin for 20 min at room temperature, diluted 10-fold with N-free medium, centrifuged, and resuspended in fresh N-free medium containing 0.01% chicken egg white trypsin inhibitor. For immunoblotting, the remaining cells were washed twice more with N-free media containing 0.01% chicken egg white trypsin inhibitor before analysis by SDS-PAGE and immunoblotting (Misamore et al., 2003).

Generating Chlamydomonas Cells Expressing Mutant Forms HAP2-HA

ΔTRA$_{184\text{-}6}$, R$_{185}$A, R$_{185}$K, R$_{185}$Q, RA$_{185\text{-}6}$AR and FW$_{192\text{-}3}$AA mutant forms of *Chlamydomonas* HAP2 (GI:288563868) were performed using standard PCR methods. The PCR fragments were inserted into the HAP2-HA plasmid at BglII/*Nru*I sites using In-fusion Dry-down PCR Cloning Kit (Clontech). All plasmids were confirmed by DNA sequencing. Plasmids were transformed into the *40D4 hap2* mutant using electroporation and selected on paromomycin plates (Liu et al., 2015). Colonies were selected based on PCR identification of the transgene and confirmation of HAP2-HA expression by immunoblotting (Liu et al., 2008).

Expression and Purification of the Soluble HAP2 Ectodomain

Codon-optimized synthetic cDNA corresponding to a soluble C-terminally truncated version of the HAP2 ectodomain (HAP2e) comprising residues 23-592 from *Chlamydomonas reinhardtii* was cloned into a modified *Drosophila* S2 expression vector described previously and transfection was performed as reported earlier (Krey et al., 2010). For large-scale production, cells were induced with 4 μM $CdCl_2$ at a density of approximately 7x10^6 cells/ml for 8 days, pelleted, and the soluble ectodomain was purified by affinity chromatography from the supernatant using a StrepTactin Superflow column followed by size exclusion chromatography using a Superdex200 column in 10mM bicine pH9.3. Pure protein was concentrated to approximately 20 mg/ml.

SEC-MALLS

Purified HAP2e ectodomain was subjected to SEC using a Superdex 200 column (GE HealthCare) equilibrated with the indicated buffers. Separations were performed at 20°C with a flow rate of 0.5 mL min^{-1}. Online MALLS detection was performed with a DAWN-HELEOS II detector (Wyatt Technology, Santa Barbara, CA, USA) using a laser emitting at 690 nm. Online differential refractive index measurement was performed with an Optilab T-rEX detector (Wyatt Technology). Data were analyzed, and

weight-averaged molecular masses (Mw) and mass distributions (polydispersity) for each sample were calculated using the ASTRA software (Wyatt Technology).

Production of Monoclonal Antibodies Targeting *C. reinhardtii* HAP2

BALB/c mice were immunized subcutaneously with 10 μg of recombinant HAP2e in Complete Freund's adjuvant and boosted five times with the same antigen dose in Incomplete Freund's adjuvant. Mice splenocytes were fused to P3U1 myeloma cells and growing hybridomas were selected in an ELISA test on plates coated with 1 μg/ml HAP2e. Specific IgG producing hybridomas were further subcloned by limiting dilution. Monoclonal antibody K3 was purified using Protein G HiTrap columns (GE Healthcare) according to the manufacturer's instructions followed by SEC in PBS using a Sdx200 column.

Liposome Co-flotation Experiments

DOPE (1,2-dioleoyl-sn-glycero-3-phosphoethanolamine), DOPC (1,2-dioleoyl- sn-glycero-3-phosphocholine), cholesterol and sphingomyelin were purchased from Avanti Polar Lipids. Liposomes were freshly prepared by the freeze–thaw and extrusion method using molar ratios of 1/1/3/1 of DOPE/DOPC/cholesterol/sphingomyelin. 0.7 μM purified HAP2e were mixed with 8 mM liposomes and incubated 1h at 25°C in 100 μL PBS. Samples were then adjusted to a final concentration of 20% sucrose, overlayed with a 5%–60% sucrose gradient (in PBS) and centrifuged overnight at 4°C at ~152.000 x g. Fractions from the top, middle and bottom of the gradient were analyzed by immunoblotting using specific anti-HAP2 monoclonal antibodies and the bands quantified using the GeneTools Syngene software. The percentage of HAP2e in either fraction was calculated as the ratio between HAP2e in individual fractions and total HAP2e (sum of HAP2e in top and bottom fractions).

Electron Microscopy

Purified HAP2e (*C. reinhardtii*) mixed with liposomes was spotted on glow discharged carbon grids (CF300, EMS, USA), negatively stained with 2% phosphotungstic acid (PTA) pH 7.4, analyzed with a Tecnai G2 Bio-Twin electron microscope (FEI, USA) and imaged with an Eagle camera (FEI, USA). For cryo-electron microscopy liposomes alone or liposomes mixed with purified HAP2e were applied on a glow discharged Lacey Carbon grid (Agar Scientific, UK). Samples were plunge-frozen in liquid ethane using an automated system (Leica EMGP, Austria) and visualized on a Tecnai F20 electron microscope operating at a voltage of 200 kV. Image frames were recorded in low-dose mode on a Falcon II direct electron detector (FEI, USA).

Native Gel Analysis

Purified recombinant HAP2e in PBS and HAP2e from the top fraction of the sucrose gradient solubilized with 4% CHAPS were independently purified by size-exclusion chromatography using a Superdex 200 Increase column. Elution fractions corresponding to the respective proteins were further analyzed on a 4%–16% native gradient gel using the nativePAGE Novex Bis-Tris gel system (Invitrogen) followed by silver staining.

Crystallization and Structure Determination

Crystals of HAP2e were obtained using in situ proteolysis as described before (Dong et al., 2007). Briefly, subtilisin dissolved in 10mM Tris pH8, 30mM NaCl at 10mg/mL was added to protein solution (12-14 mg/ml in 10 mM bicine, pH 9.3) on ice immediately prior to crystallization trials in a 1/100w:w ratio. Crystals of HAP2e were grown at 293K using the hanging-drop vapor-diffusion method in drops containing 1 μL protein/protease solution mixed with 1 μL reservoir solution containing 100 mM HEPES pH7.5, 2% 2-Propanol, 100 mM sodium acetate, and 12%–14%w/v PEG 8000. Diffraction quality rod-like crystals appeared after 1 week and were flash-frozen in mother liquor containing 30% (v/v) MPD.

Data collection was carried out at the Swiss Light Source (PX I), the ESRF (ID30A-3), and the Synchrotron Soleil (Proxima1). Data were processed, scaled and reduced with XDS (Kabsch, 1988), Pointless (Evans, 2006) and programs from the CCP4 suite (Collaborative Computational Project, 1994). A single-wavelength anomalous dispersion (SAD) dataset was collected from a single crystal of HAP2e from *C. reinhardtii* soaked for 6 hr in 2 mM K2PtCl4 solution in cryo buffer. Data were collected at the LIII edge of Platinum (1.072 Å) on a single crystal using low-dose (0.5 MGy per 360°), high-redundancy (5 × 360°) fine φ-sliced collection strategy using five crystal orientations by means of a high-precision multi-axis PRIGo goniometer (Weinert et al., 2014). An initial set of experimental phases was obtained by the Single Isomorphous Replacement method using autoSHARP (Vonrhein et al., 2007) with the Platinum derivative and a highly isomorphous native dataset. Starting phases were improved by consecutive cycles of manual building and combination with phases derived from molecular replacement using the partial model as search model in Phaser (MR-SAD) (McCoy et al., 2007). After building an initial poly-alanine model accounting for ~50% of the Cα atoms these phases were further refined using the anomalous signal of a highly redundant Sulfur-SAD dataset collected at a wavelength of 2.06641 Å on crystals of the native protein following a similar collection strategy as mentioned above (Weinert et al., 2014). Model building was performed using Coot (Emsley et al., 2010), and refinement was done using AutoBuster (Bricogne et al., 2010) with repeated validation using MolProbity (Chen et al., 2010). The final model includes amino acids 24 to 581 (see linear diagram in Figure 1B), with internal breaks at loops 69-97, 152-156, 167-182, 194-205, 238-283 and 330-345, corresponding to disordered loops that are marked with a gray background in the *C. rheinhardtii* sequence in Figure 5D (top sequence) and as dashed tubes in the ribbon diagrams (Figures 4 and 5A–5C). Clear electron density was observed for one N-linked and one O-linked glycan chain (attached to N_{497} and T_{577} in domain III).

QUANTIFICATION AND STATISTICAL ANALYSIS

Data are presented as mean ± SD unless otherwise indicated in figure legends and experimental repeats are indicated in figure legends.

DATA AND SOFTWARE AVAILABILITY

The accession number for the atomic coordinates and structure factors reported in this paper is PDB: 5MF1.

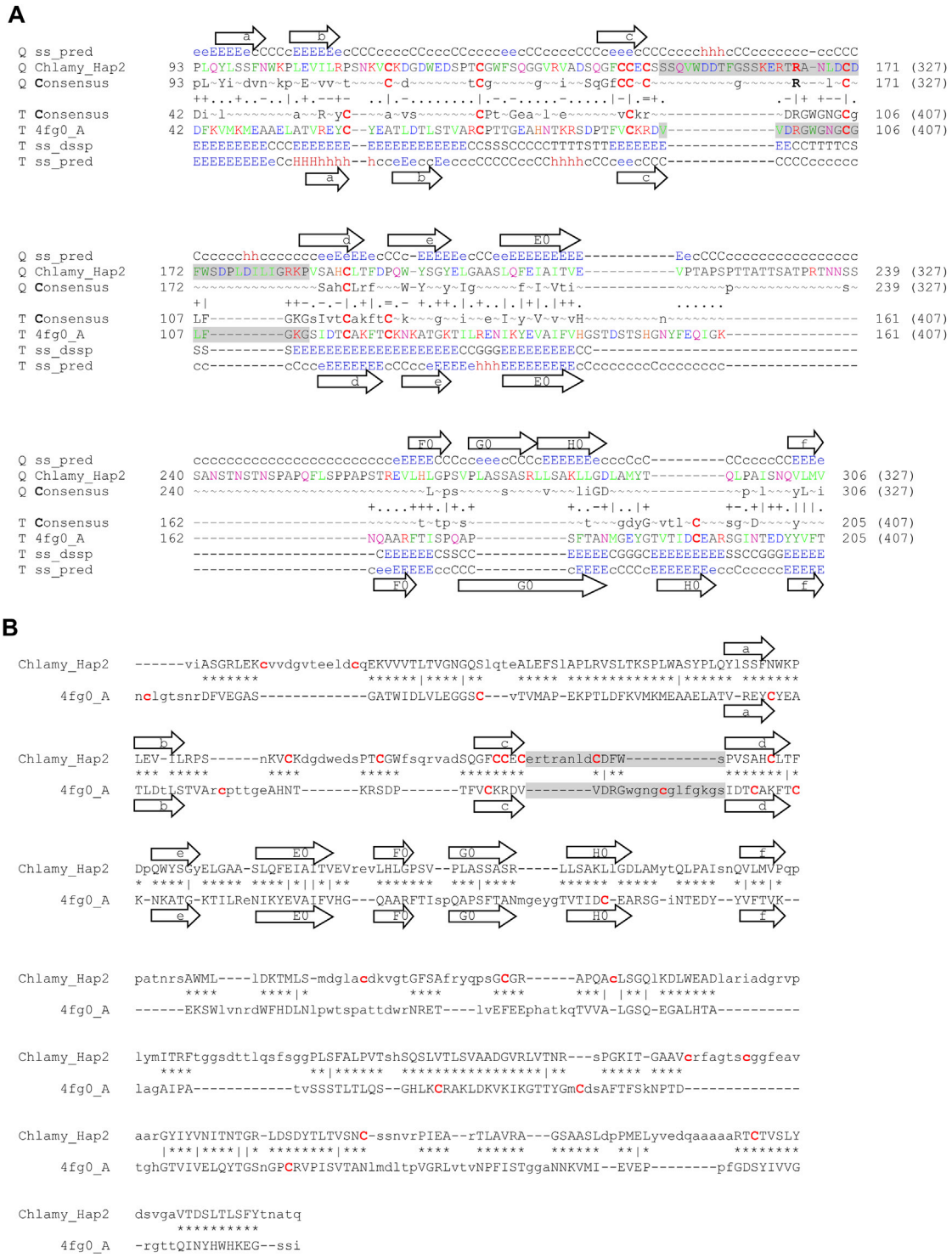


Figure S1. Identification of a HAP2 Region Homologous to Class II Fusion Proteins, Related to Figure 1.

(A) HHpred alignment between *Chlamydomonas* HAP2 N-terminal region and St. Louis Encephalitis Virus E (SLEV E) protein (PDB: 4fg0, chain A; Luca et al., 2013). The fusion loop region of SLEV E and the corresponding aligned region of *Chlamydomonas* HAP2 are highlighted in gray. "Q ss_pred," "Q Chlamy_Hap2," and "Q Consensus" lines show predicted secondary structure, amino acid sequence and profile consensus sequence of *Chlamydomonas* HAP2. "T Consensus," "T 4fg0_A," "T ss_dssp" and "T ss_pred" lines show profile consensus sequence, amino acid sequence, observed secondary structure in the PDB file using DSSP, and predicted secondary structure of SLEV E. Secondary structure elements taken from the *Chlamydomonas* HAP2 and SLEV E structures (by DSSP) are shown as arrows and named according to Figure 5 in the main manuscript.

(legend continued on next page)

(B) DaliLite structural alignment of the *Chlamydomonas* HAP2 and SLEV E crystal structure (pdb: 4fg0, chain A). Structurally aligned residues are in capital letters and indicated by "*" (different residue type) or "|" (same residue type). The fusion loop region of SLEV E and the corresponding aligned region of *Chlamydomonas* HAP2 are highlighted in gray. The coordinates of *Chlamydomonas* HAP2 in this region lack the amino acid segments "SSSQVWDDTFGSSK" between "CEC" and "ert" and "DPLDILIGRK" between "FWs" and "PVS." The secondary structural elements covered in the HHpred alignment are shown as arrows in this DaliLite alignment.

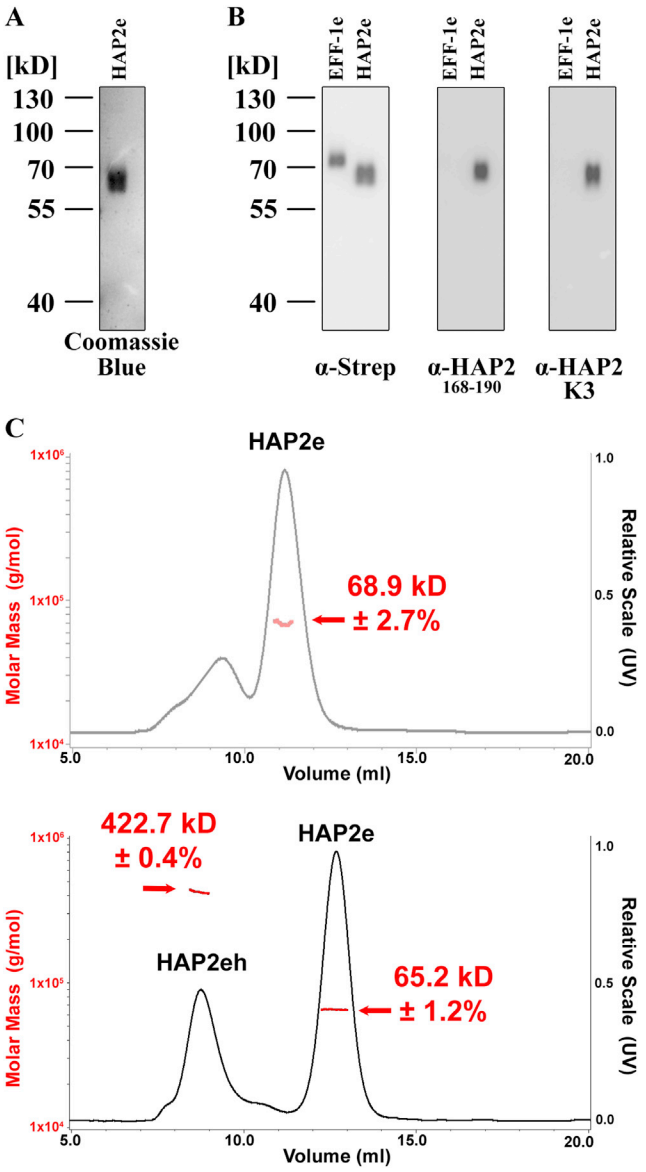


Figure S2. Purification of HAP2e and Specificity of HAP2 Antibodies, Related to Figure 2

(A) HAP2e purity was assessed by SDS-PAGE followed by Coomassie blue staining.

(B) The specificity of the α-HAP2$_{168-190}$ polyclonal antibody (middle) and the α-HAP2 monoclonal antibody K3 (see STAR Methods section) was analyzed by immunoblotting of HAP2e and a soluble ectodomain of the *C. elegans* fusion protein EFF-1 (EFF-1e).

(C) SEC analysis of HAP2e on a Superdex 200 column revealed the presence of two peaks, one corresponding to a monomer (right peak) and one corresponding to a higher molecular weight oligomer (left peak). Under low ionic strength conditions at pH 9.3 (here shown in the top panel without NaCl in Bicine buffer 10mM) the amount of multimerized protein was less than in PBS, at higher ionic strength conditions and pH 7.4 (lower panel). SEC-MALLS analysis indicated that the peak to the right corresponds to a monomer, and the peak to the left in the PBS run to a hexamer. The y axis on the right gives the UV absorption at 280 nm, on a relative scale from 0 to 1 (plotted in black) while the y axis on the left gives the molecular mass by MALLS (plotted and labeled in red).

In retrospect, our interpretation for the oligomers is that with time, recombinant HAP2e spontaneously trimerizes into the post-fusion form, and that these trimers associate into rosettes to mask the exposed hydrophobic region of the fusion loops. The fact that this species increases with ionic strength supports this interpretation. The hexameric form is likely formed of two such trimers facing each other via the fusion loops region (the simplest rosette), but this remains to be verified.

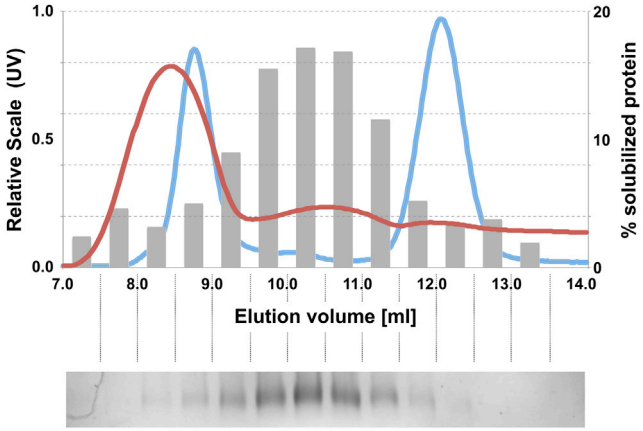


Figure S3. Lipid Insertion of HAP2, Related to Figure 3

Recombinant *C. reinhardtii* HAP2 was subjected to SEC on a Superdex 200 column and the fraction eluting as monomer (HAP2e, peak on the right in the blue curve) was mixed with liposomes and loaded onto a sucrose gradient (as in Figure 3E). Upon centrifugation (see STAR Methods), the fraction recovered from the top of the gradient was again run in the same SEC column (red plot). This time the elution profile showed a major peak of UV 280 nm absorption near the void volume, likely resulting from the UV absorption of large mixed detergent/lipid micelles, as SDS-PAGE analysis of the individual elution fractions of the SEC column followed by Coomassie staining (shown below the profile) revealed that this peak contained no protein. The protein bands in the other fractions were quantified and the amount of HAP2e in the individual fractions is shown in the superposed histogram, indicating that solubilized HAP2e elutes from the column in between the peaks corresponding to the monomeric and the hexameric HAP2e fraction, as expected for a trimer.

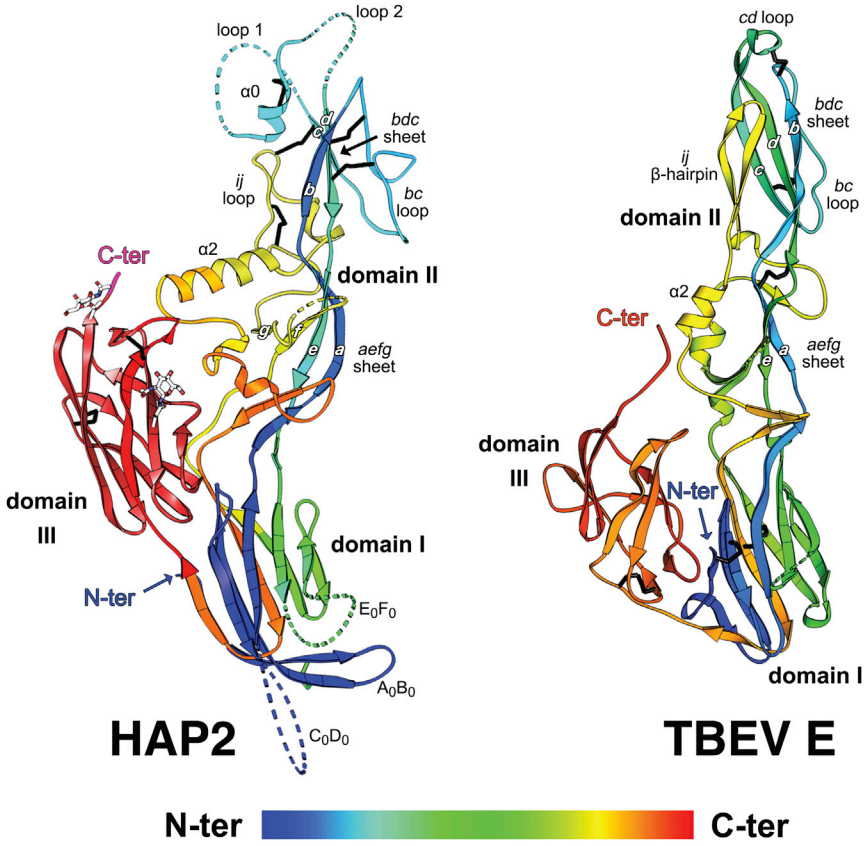


Figure S4. Class II Topology of HAP2, Related to Figure 5

Ribbon representation of a trimer subunit of *C. reinhardtii* HAP2 (left) and the flavivirus TBEV E protein (PDB 1URZ; right), ramp-colored from N to C terminus according to the color ramp in the bar underneath, to highlight the secondary and tertiary structure matches throughout the length of the polypeptide chain. Disulfides are drawn as black sticks. Salient secondary and tertiary structure motifs discussed in the text are labeled.

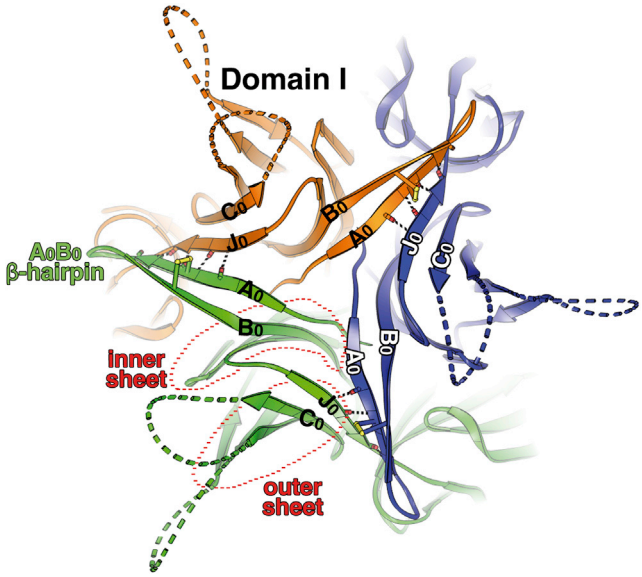


Figure S5. Stability of the HAP2 Trimer, Related to Figure 4

The *C. reinhardtii* HAP2 trimer seen down the 3-fold axis, from the side opposite to the membrane-interacting end, colored according to subunits. The long A_0B_0 β-hairpin augments the outer β sheet of the adjacent subunit, with strand A_0 running parallel to J_0 of the adjacent clockwise subunit. Inter-subunit main chain β-interactions are depicted as black dotted lines. Red dotted contours outline the outer and inner sheets of the subunit in green. The disulfide bonds connecting β strands A_0 and B_0 are shown in yellow.