X-ray structures of the post-fusion 6-helix bundle of the human syncytins and their functional implications

Katinka Ruigrok, Marie-Christine Vaney, Julian Buchrieser, Eduard Baquero, Jan Hellert, Bruno Baron, Patrick England, Olivier Schwartz, Felix Rey, Marija Backovic

X-ray structures of the post-fusion 6-helix bundle of the human syncytins and their functional implications

Katinka Ruigrok[a,b], Marie-Christine Vaney[a], Julian Buchrieser[d], Eduard Baquero[a], Jan Hellert[a], Bruno Baron[c], Patrick England[c], Olivier Schwartz[d,e], Felix A. Rey[a,#], Marija Backovic[a,#]

[a] Unité de Virologie Structurale, Institut Pasteur, Département de Virologie, 28 rue du Dr Roux, 75015 Paris, France; CNRS, UMR3569, Paris, France.

[b] Université Paris Descartes, Sorbonne Paris Cité, Paris, France.

[c] Plate-Forme de Biophysique des Macromolecules et de leurs Interactions, Institut Pasteur, 25 rue du Docteur Roux, 75015 Paris, France.

[d] Unité de Virus et Immunité, Institut Pasteur, Département de Virologie, 28 rue du Dr Roux, 75015 Paris, France; CNRS, UMR3569, Paris, France.

[e] Vaccine Research Institute, Creteil, France.

[#]Corresponding authors: Marija Backovic (marija@pasteur.fr), Felix Rey (rey@pasteur.fr)

Institut Pasteur, 28 rue du Dr Roux, 75015 Paris, France

Tel +33 1 40 61 31 29

ABSTRACT

The retroviral-envelope-derived proteins syncytin-1 and -2 (syn1 and syn2) drive placentation in humans by forming a syncytiotophoblast - a structure allowing for an exchange interface between maternal and fetal blood during pregnancy. Despite their essential role, little is known about the molecular mechanism underlying the syncytins' function. We report here the X-ray structures of the syn1 and syn2 transmembrane subunit ectodomains, featuring a 6-helix bundle (6HB) typical of the post-fusion state of gamma-retrovirus and filovirus fusion proteins. Contrary to the filoviruses, for which the fusion glycoprotein was crystallized both in the post-fusion and in the spring-loaded pre-fusion form, the highly unstable nature of the syncytins' pre-fusion form has precluded structural studies. We undertook a proline-scanning approach searching for regions in the syn1 6HB central helix that tolerate the introduction of helix-breaker residues and still fold correctly in the pre-fusion form. We found that there is indeed such a region, located two α-helical turns downstream a stutter in the central coiled-coil helix - precisely where the breaks of the spring-loaded helix of the filoviruses map. These mutants were fusion-inactive as they cannot form the 6HB, similar to the "SOSIP" mutant of HIV Env that allowed the high-resolution structural characterization of its labile pre-fusion form. These results now open a new window of opportunity to engineer more stable variants of the elusive pre-fusion trimer of the syncytins and other gamma-retroviruses envelope proteins for structural characterization.

1. INTRODUCTION *

Syncytins are cell-cell fusion proteins that drive placenta formation by inducing the merger of peripheral blastocyst cells, called cytotrophoblasts, into a giant syncytium called syncytiotrophoblast. This structure provides an exchange interface between maternal and fetal blood, thereby establishing a nutrient supply to the fetus [1]. Despite the syncytins' key role in placental morphogenesis and likely in a number of pregnancy diseases [2], as well as their involvement in osteoclast fusion during bone formation [3], they remain poorly characterized molecularly. The syncytin genes are among the numerous endogenous retroviral elements [4] that account for 8% of the human genome [5]. While the majority of such elements have accumulated mutations and sequence rearrangements that make them non-functional, some have preserved their open reading frames due to selective advantages for the host. The syncytins are prime examples of such appropriations, involving domestication of the retroviral genes coding for the viral envelope protein (Env), which mediates fusion of the viral lipid envelope with the plasma membrane of target cells during entry.

One or more syncytins have been found in all so far investigated placental mammals, including marsupials [2, 6], and also in more distant vertebrates such as viviparous lizards [7]. They have been the product of independent endogenization events, with the earliest syncytin capture envisioned as the driving force behind the evolution of egg-laying vertebrates into placental mammals [1]. In humans, there are two syncytins, 1 and 2 (syn1 and syn2), that derive from the *env* genes of the human endogenous retroviruses W and FRD (HERV-W and HERV-FRD), respectively [8]. Both

* ABBREVIATIONS : **6HB** – six-helix bundle, **aa** – amino acid; **BLV** – bovine leukemia virus; **EBOV** – Ebola virus; **Env** – envelope protein; **ER** – endoplasmic reticulum; **GFP** – green fluorescent protein; **HERV** – human endogenous retrovirus; **HIV** – human immunodeficiency virus; **HR** – heptad repeat; **HTLV** – human T-lymphotropic virus; **MARV** – Marburg virus; **MPER** – membrane-proximal external region ; **MPMV** - Mason-Pfizer monkey virus; **PDB** – Protein Data Bank; **SEC** – size exclusion chromatography; **SEM** - standard error of the mean value; **SP** – signal peptide; **SU** – surface subunit; **Syn1** – human syncytin-1 ; **Syn2** – human syncytin-2; **TEV** - tobacco etch virus; **TM** – transmembrane subunit; **TM^A** – transmembrane anchor; **TM^E** – ectodomain of the transmembrane subunit; **WT** – wild type.

are gamma-retroviruses [9] that entered the human genome around 25 and 45 million years ago, respectively [10-12]. Human syncytins possess a gamma-type Env protein, which is present not only in gamma-retroviruses but also in delta-, alpha- and in a subset of the beta-retroviruses [13].

The Env protein of all retroviruses infecting vertebrates is a class I viral fusion protein, a class that also includes the fusion glycoproteins of the otherwise unrelated influenza viruses, coronaviruses, paramyxoviruses, pneumoviruses, filoviruses and arenaviruses [14]. Class I proteins are type 1 single-pass transmembrane proteins that fold in the ER of the infected cell as a precursor that in most cases trimerizes directly upon folding. An activating proteolytic step cleaves the precursor in two - an N-terminal portion, termed "surface" (SU) subunit in retroviruses - which in general has a receptor binding function, and a C-terminal subunit, which carries the membrane fusion function and is termed "trans-membrane" (TM) subunit, as it includes the trans-membrane anchor near its C-terminal end (Fig. 1A). The SU subunit remains peripherally associated, and in some cases is covalently linked to TM by a disulfide bond [13]. Cleavage of the precursor into SU and TM traps the complex in a metastable state (the "pre-fusion" form), with the TM subunit spring-loaded underneath a crown made by SU. Interactions with target cells induce the release of the SU crown, allowing TM to spring out and undergo a major conformational change that drives membrane fusion [15]. A specific feature of the gamma-type Env proteins is the presence of two cysteine-rich motifs in the primary sequence, CXXC in SU and CX_6CC in TM (where X is any aa), which control formation of an inter-subunit disulfide bond in the pre-fusion form that is reduced upon interaction with a receptor to release SU and induce the fusogenic conformational change in TM required for entry [16].

An important functional element of the TM subunit is the "fusion peptide", a hydrophobic segment located at its N-terminal end and normally buried under the SU crown. The fusogenic conformational change entails first projection of the fusion

peptide to a distance about 10 nm or more away from its original location, such that it can reach and insert into the target cell membrane. In this transient conformation, the TM protein has the N- and C-termini at opposite ends of a long trimeric rod, thereby bridging viral and cellular membranes at a distance of about 15 nm [17]. The trimer protomers bend in half, adopting a "hairpin" conformation that brings the N- and C-terminal ends of the TM ectodomain into contact, forcing apposition of the two membranes at a distance within 1nm. The conformational change of TM thus lowers a kinetic barrier arising from short-range repulsive forces between the two membranes - resulting from the necessary dehydration of both outer leaflets to allow inter-membrane lipid contacts [18] - to catalyze lipid merger.

The X-ray structure of the TM ectodomain (TM^E) in the post-fusion hairpin conformation is known for Env from several retroviruses [19-25]. These structures showed a long α-helix (termed N-helix) immediately downstream of the fusion peptide, which forms a characteristic central, parallel trimeric α-helical coiled-coil. The amino acid sequence in this region has characteristic heptad repeats (HR), represented as positions *abcdefg* (Fig. 1B), where *a* and *d* are typically occupied by apolar amino acids that form the hydrophobic core of the coiled coil. Downstream the N-helix, the polypeptide chain makes an overall 180 degrees turn, such that the C-terminal segment runs backward, antiparallel to the coiled-coil and inserting an α-helix (the "C-helix") along the grooves in-between the long N-helices in the trimer to complete the hairpin. The resulting 6-helix bundle (6HB) thus has the C-terminal segment connecting to the trans-membrane anchors in proximity of the fusion peptides projecting from the N-terminal end. The structure of the pre-fusion form of the retroviral Env ectodomain, in which the N-helix is broken into several short helices to maintain a more compact form underneath the SU crown, remains a challenge. HIV-1, which has a "beta-type" Env, is the only retrovirus for which pre-fusion Env structure is known (reviewed in [26] and [27]). In this case, as with other class I fusion proteins, the introduction of mutations

interfering with formation of the post-fusion 6HB were necessary to stabilize the pre-fusion trimer and make it accessible to structural studies [28].

Structural studies on syncytins have been limited to the X-ray structure of a syncytin-2 TM^E fragment missing around half of the TM^E [29]. We present here the structures of the complete syncytin-2 and syncytin-1 TM^E in post-fusion conformation, which allowed us to perform structural comparisons of syncytins with other class I fusion proteins, confirming the anticipated structural homology to the retroviral gamma-type Envs, but also to GP2 fusion subunit of filoviruses and reptarenaviruses. Our results identified a vertical string of 9 salt bridges crisscrossing the N-terminal part of the N-helix and the C-terminal part of the C-helix, found in syn1 but not in syn2, in line with the higher thermal stability measured for syn1-TM^E. The new structures, together with their close structural homology with the filovirus counterparts, guided further functional studies with full-length syn1 variants carrying helix-breaker proline substitutions along the N-helix region. We thus identified locations in which the proline appeared to block formation of the post-fusion form - as monitored by their loss of function in a syncytium formation assay - but remained compatible with folding in the spring-loaded pre-fusion form as they were trafficked to the cell surface as efficiently as wild type protein.

2. RESULTS

2.1. Construct design and structure determination

Although syn1 and syn2 have the same number of residues in total (538 aa) and their TM^E are also about the same length (~140 aa), syn1 has a shorter SU and a longer cytosolic tail (Fig. 1A). We used structural information available for a number of post-fusion retroviral TM^E conformations [22, 24, 25] together with aa sequence alignments and secondary structure / hydrophobicity plots [30, 31] to define the boundaries of the constructs we used for the structural studies. These analyses (Fig. S1) suggested that the fusion peptide spans the TM N-terminal 23-24 aa, and that the HRs begin 6-7 residues downstream, encompassing around 48 aa in total for both syncytins. The N-terminal end in the previously reported X-ray structure of a syn2-TM^E fragment [29] was about 10 residues downstream these predictions, suggesting that the N-helix was truncated. In the C-terminal segment, our predictions indicated a roughly 25-residue long α-helix downstream the characteristic gamma-type Env CX_6CC motif, which would correspond to a C-helix that was absent in the reported syn2-TM^E structure (Fig. 1A, Fig. 2; PDB accession code **1Y4M** [29]). A relatively hydrophobic element is present between the predicted C-helix and the trans-membrane anchor (TM^A), most likely corresponding to the amphipathic membrane proximal external region (MPER) identified in a number of other retroviral Env proteins. From these analyses, we designed expression vectors for a recombinant TM^E core, termed here syn1-TM^E (residues 342-435) and syn2-TM^E (residues 375-468) starting immediately downstream of the fusion peptide and ending before the putative MPER (Fig. 1B).

We cloned these constructs into a pET28 vector for production in *E. coli*, including a histidine tag fused at the N-terminus. Upon purification by nickel affinity followed by size exclusion chromatography (SEC), we obtained yields of about 30 milligrams of each protein per liter of cell culture. Both proteins formed trimers with a

molecular mass of ~36 kDa as judged from their SEC elution profiles (data not shown). Syn1-TM^E crystallized in a rhombohedral space group (H32) and the crystals, which contained one protomer in the asymmetric unit, diffracted to 2.1Å resolution. Syn2-TM^E formed trigonal crystals (space group $P3_221$) that diffracted to 2.2Å and had a trimer in the asymmetric unit. The corresponding structures were determined by molecular replacement using the truncated syn2-TM^E structure as search model (PDB accession number **1Y4M**). The final map displayed clear electron density for residues 345-433 in the crystals of syn1-TM^E, with the first three and last two residues of the construct being disordered. For syn2-TM^E, the trimer is local and residues 380-465 are resolved in all three chains of the trimer, with individual protomers displaying additional N- or C-terminal extensions ordered because of stabilization by crystal packing contacts. The atomic models were refined to the maximum resolution of the diffraction data to R_{work}/R_{free} of 0.22/0.24 and 0.21/0.26 for syn1- and syn2-TM^E, respectively. The statistics concerning the structure determination and refinement of the atomic model are listed in Table S1. The crystallographic details are described in the Materials and Methods section.

2.2. The syncytins are typical gamma-type Env proteins

The structures showed that the syn1 and syn2 6HBs are formed by an N-helix of 12 turns (about 44 aa organized in 6 HRs) followed by a "turn" (T) region (21 aa), where the chain reverses orientation, and a 6 turns C-helix (23 aa) (Fig. 2). Aligning the 6HBs of syn1 and syn2 resulted in 0.9Å root mean square deviation (rmsd) for 258 aligned Cα atoms (Table S2). A DALI search [32] using the syn1- and syn2-TM^E structures as queries, showed that they are most similar to the 6HB of the beta-retrovirus Mason-Pfizer monkey virus (MPMV), which has a gamma-type Env [33], and also to the prototypical gamma-retrovirus xenotropic murine leukemia virus-related virus (XMRV) (Fig. 2, lower panel). The observed structural similarity is in line with the

51% and 45% aa sequence identity of the MPMV TM^E core with syn1 and syn2, respectively, which is higher than between syn1 and syn2 (44% aa sequence identity). In three dimensions (3D), the 6HB of syn1 and syn2 align with 0.9Å and 1.2Å rmsd, respectively, onto the MPMV counterpart, with 258 aligned Cα atoms (Table S2). The XMRV 6HB is slightly more distant, but still highly similar, as are the 6HBs of the delta-retroviruses human T-lymphotropic virus (HTLV) and bovine leukemia virus (BLV) Env proteins, which are also gamma-type Envs. The most obvious difference is that the BLV and HTLV proteins have two or a single short helix, respectively, replacing the C-helix in the 6HB [24, 25] (Fig. S2A). The corresponding rmsd values upon 3D alignment on the syn1 and syn2 6HBs are given in Table S2 and Fig. S2B.

Polar interactions stabilizing the 6HB. As also observed in the previously reported gamma-type TM^E structures [22, 24, 25], the 6HB of the syncytins features two "belts", or networks of hydrogen bonds: belt 1 ties the N-helix at turn 7 to its adjacent protomers in the trimer, and belt 2 at turns 9 and 10, close to the bottom of the bundle (Fig. 3A). Belt 1 involves the side chains of several glutamine residues in syn1 and syn2. In the 6HB of MPMV and XMRV TM^E, numerous electrostatic interactions were described between charged residues mapping to the same region (Fig. 3D) [22]. In addition, the syn1 6HB contains a string of 8 inter-chain and 1 intra-chain salt bridges that tie the first 5 turns of the N-helix to the C-helix of the adjacent protomer in a vertical fashion, i.e. by establishing numerous successive inter-chain and intra-chain contacts (Fig. 3B). The corresponding residues in the C helix of the syn2 6HB are uncharged (Fig. S2, Fig. 3D), and therefore no salt bridge is formed. Instead, the syn2 6HB features two different salt bridges, one inter- and one intra-chain, in the upper half of the bundle (Fig. 3C, Table S3). Akin to what we observe here in syn2, electrostatic interactions also clip the termini of the N- and C-helices of MPMV and XMRV at the open end of the hairpin, although they are more dispersed and less numerous (1 and 2, respectively) (Fig. 3D).

The ionic character of the interactions stabilizing the syn1 6HB is reflected in the negatively charged electrostatic surface potential of the upper half of the coiled-coil in syn1, which is uncharged in syn2 (Fig. 4A). To understand the contribution of these differences to the stability of the 6HB, we measured the melting temperature (T_m) of the syn1 and syn2 6HBs using differential scanning fluorimetry. We found that the syn-1 6HB had a T_m of 85°C, about 20°C higher than that of the syn2 6HB (Figure 4B), highlighting the higher stability of the former.

At the core of belt 2 there is a conserved QNRR motif (aa 381-384 in syn1 and 415-418 in syn2), with an asparagine in place of the hydrophobic residue expected at position *a* or *d* in the HR pattern (Fig. 3D). The asparagine side chain coordinates a central chloride ion on the 3-fold molecular axis, as seen previously in the HTLV-1, BLV, MPMV and XMRV TM^E structures. We refer to the interacting residues in the QNRR motif according to their location relative to the conserved asparagine (N^0), as Q^{-1}, R^{+1}, R^{+2}. In this scheme, the conserved salt bridge connects R^{+2} with E^{+20}. Both residues were shown to be essential for the fusion activity of XMRV, MPMV and HTLV Envs [22, 34]. The TM^E of the delta-retroviruses contains a second, upstream hydrophilic layer, which is in the case of HTLV-1 gp21 organized around Asn364 that traps a second chloride ion, and in BLV gp21 around Thr353 that chelates a water molecule on the 3-fold axis. These additional hydrophilic layers of inter-chain polar contacts established around a second trapped molecule / ion map to the same region as belt 1 in XMRV, MPMV and the syncytins 6HB (Fig. S2A).

The polar interactions in belt 2, with the asparagine side chain pointing into the coiled-coil cavity, are likely to be important for establishing the correct register and orientation of the helices, as has been demonstrated for other coiled-coils [35, 36]. The interactions between the residues of the N- and C- helices of the same or the adjacent protomer in the trimer (Fig. 3A (belt 1), 3B and 3C) might play a role in stabilization of the 6HB, which is in turn important to drive fusion. A direct test of the role of the

residues involved in these interactions by site-directed mutagenesis is complicated by the fact that the 6HB forms only after the fusogenic conformational change, and the lack of functionality of any given point mutant could also be due to an alternative interaction in the pre-fusion form. A detailed analysis of these residues therefore awaits the knowledge of the structure of the TM^E moiety in the context of a pre-fusion syncytin trimer.

The immunosuppressive (ISU) motif in syn2. The T loop at the chain reversal region starts with a 3/10 helical turn, termed η_T helix, followed by a short α-helix (T helix, ~2 turns; Fig. 2). The single residue responsible for the ISU properties of syn2 (Gln427) resides in the η_T helix (Fig. 1B and 2). Syn1, which does not have ISU properties, has R393 at the corresponding position. Another residue located in the syn2-TM^E T loop, A433, is important for its fusion activity, and only the simultaneous introduction of these two mutations into syn1 i.e. R393Q and F399A (Fig. 1B), converts syn1 into a functional protein with both ISU and fusogenic activities [37]. These pairs of residues, Q427 and A433 in syn2-TM^E, are conserved in MPMV (Q471, A477), while a glutamic acid replaces the glutamine in XMRV (E526) (Fig. 2). Even though syn1 is not immunosuppressive, the main chain conformation in the region spanning the ISU is the same as in syn2, MPMV and XMRV, which all have functional ISU properties.

2.3. Similarities to non-retroviral 6HBs

In addition to the expected structural similarity to the gamma-retrovirus type 6HB, syn1 and syn2-TM^E are also structurally very close to the fusion proteins of filoviruses such as Ebola virus (EBOV) and Marburg virus (MARV), as well as the filovirus-like fusion protein of reptarenaviruses such as the California Academy of Sciences virus (CASV), as already pointed out [38, 39]. This is reflected in the low rmsd values obtained upon superposition of their trimers (≤1.9Å for ≥83% residues aligned;

Table S2), as well as conservation of structural motifs such as CX_6CC, ISU, [40], and belts 1 and 2 (Fig. S2A). Such a close conservation of features is not observed when comparing syncytins to the 6HB of retroviruses bearing a beta-type Env protein, like the lentiviruses, which lack the conserved $CXXC/CX_6CC$ motifs and do not form the inter-subunit TM-SU disulfide bond.

In the filoviral-like proteins the length of the structural elements of the 6HB is similar to the gamma-retroviral counterparts, with the C-helices showing some variability (Table S2, Fig. S2B). The QNRR-E^{+20} motif at the core of belt 2 is LNRK-D^{+21} and INRH-D^{+21} in EBOV and MARV GP2, respectively (Fig. S2A). The chloride ion coordinated by the central asparagine is in addition coordinated by a serine side chain located three residues upstream – at the *a* position of the same HR [41, 42] (Fig. S2). CASV GP2 is slightly different, as the central asparagine is replaced by a histidine residue that coordinates a central water molecule on the trimer axis [39]. The high structural similarity between the syncytins 6HB and the filovirus counterpart is of particular interest, as the fusion proteins of EBOV and MARV have also been structurally characterized in their pre-fusion form [43-45].

Potential break points of the N-helix in the pre-fusion form. The X-ray structures of a number of class I fusion proteins from different viral families have been determined in both pre- and post-fusion forms (reviewed in [14]), and the breaks in the N-helix in the spring-loaded form have been mapped. In several of these structures, including filoviral GP, the break points were found downstream of a special turn in the N-helix that makes a "stutter" in the central coiled coil. Stutters are defined as HR disruptions in the form of 4-residue insertions in between two HRs [46], and are found in most alpha-helical coiled-coils. The stutter-induced deviation from the ideal geometry results in local unwinding of the coiled-coil superhelix, and while the helical axes are almost straight for the first three turns of the N-helix, their twist becomes evident as the coil progresses

towards the C-terminus (Fig. 2). Stutters have been proposed to impart structural plasticity to regions that function as conformational switches [39], in agreement with their potential to suggest possible sites for spring-loading the N-helix. Analysis of the coiled-coil by program TWISTER [47] revealed the presence of a stutter at the 3rd turn of the N helix in both syncytin structures, residues ^{354}LNGD^{357} in syn1 and ^{388}IANN^{391} in syn2-TM^{E} (Fig. 1B, Fig. 2). A stutter is present at the same location in the 6HB of MPMV, XMRV and BLV, as well as in CASV, two helical turns upstream in HTLV and two turns downstream in the filovirus GP2 6HB (Fig. S2).

2.4. Identification of helix-breaker residues along the syncytin N-helix that do not affect transport of the full-length protein to the cell surface.

Our attempts to produce a biochemically tractable syncytin ectodomain (SU/TM^{E})$_3$ for structural studies have not been successful, and a likely reason is its instability in the spring-loaded form. We therefore explored the effect of introducing helix-breaking residues (prolines) along the whole length of the N-helix, with the goal of identifying positions that may not interfere with folding of the syncytins in their pre-fusion conformation. Such an approach was applied successfully to stabilize the pre-fusion state of HIV-1 Env [28], and later of other class I proteins, such as the spike protein of the Middle East respiratory syndrome-related coronavirus [48] or the fusion (F) protein of the human respiratory syncytial virus as well as the human metapneumovirus [49, 50].

We generated a set of single proline variants of syncytin-1, containing a helix-breaking residue at *a* and *d* positions in the N-helix: Y347P, L350P, L354P, M358P, L365P, L368P, L372P, L375P, and V379P. We monitored by flow cytometry the effect of these mutations on the surface-expression of the syn1 variants upon transient transfection (described in section 4.6.2), as indicative of a proper pre-fusion conformation. With this system, we found that the proline substitutions introduced

between turns 6 and 7 of the N-helix (L365P, L368P and L372P) led to a protein that was transported to the cell surface as efficiently as wild type (Fig. 5). In contrast, proline substitutions further upstream (turns 1, 2 and 3, Y347P, L350P, L354P, M358P) or downstream (turns 8 and 9, L375P and V379P) resulted in impaired transport to the cell surface compared to wild type (Fig. S3A and S3B, respectively), while the transfection efficiency and total syn1 expression were similar in all the samples.

As controls, we tested in parallel other sites in the 6HB, for instance the effect of altering the polar layer at the QNRR motif, a mutation that that had been tested for BLV Env [24]. We found that several substitutions at this location - syn1 N382D, N382A, and N382L - partially affected its trafficking to the cell surface (Fig. S3C), in contrast to the previously tested BLV Env variants (N367D and N367L), which were expressed at the cell surface at a level similar to that of the WT protein [25]. Mutation of the syn1 disulfide isomerization motif CXXC to SXXC in SU, (C186S), also resulted in partially impaired trafficking to the cell surface. Finally, mutating the furin cleavage site (R314S, R317S) resulted in a protein that trafficked to the cell surface without significant differences to the wild type protein (Fig. 5B).

We then used 293T cells stably transfected with either one of two complementary fragments of the green fluorescence protein (GFP) to analyze syncytia formation upon transient transfection with full-length syncytin-1 and selected variants. Syncytia formation leads to reconstitution of a functional GFP, allowing easy monitoring of cell-cell fusion by bimolecular fluorescence complementation [51] (described in detail in [52] and in the Material and Methods section 4.6.3). These cells constitutively express the syn1 receptor, the alanine-serine-cysteine transporter 2 (ASCT-2) molecule, which is required for syn1 driven cell-cell fusion [53]. As expected, the GFP signal appeared in the cells upon transfection with wild type syn1, indicative of its fusogenicity (Fig. 6, Fig. 7A). In contrast, no GFP signal was detected upon transfection with the proline variants, including those that underwent efficient trafficking to the cell surface, indicating that the

mutants were unable to induce cell-cell fusion because they could not form the required 6HB. The variants carrying N382 mutations did induce GFP signal, although at ≤20% of what was observed for the wild type syn1. N382D was more active than N382A and N382L, which showed essentially no fusion activity (Fig. 6, Fig. 7A, Fig. S4), similar to what was observed when asparagine to aspartic acid or leucine mutations were introduced into BLV TM^E [25]. Finally, the controls with the altered disulfide motif or the furin site knock-out, did not result in measurable cell fusion either, consistent with the already published data on the human syncytins [54, 55], HTLV [16], and MLV Envs [56].

3. Discussion

The structures of the syn1- and syn2-TM^E core revealed a 6HB very similar to that of the closely related gamma-type Env proteins of the gamma-retrovirus XMRV, the beta-retrovirus MPMV, and the delta-retroviruses HTLV and BLV. They also revealed close similarity to the more distant fusion glycoproteins of the filoviruses MARV and EBOV, and of the filovirus-like CASV. The two polar layers described in the structures of the XMRV and MPMV TM^E [22] are also present in the syncytins, and have their counterpart in the 6HB of all the proteins mentioned above. A number of these polar interactions on the MPMV and XMRV 6HB were tested by site-directed mutagenesis and shown to interfere with the fusion activity of the corresponding full-length protein [22]. The same residues, however, can potentially be involved in a different set of interactions in the pre-fusion form, and it is therefore difficult to draw strong conclusions about their specific role in stabilization of the post-fusion 6HB and its relation to the membrane fusogenic activity of the protein. Our studies also confirmed that the core TM^E fragment of syn2 reported back in 2005 [29] was truncated roughly at half the length of the protein core (Fig. 2), similar to the first described structure of a gamma-retroviral protein in the post fusion form, that of the murine Moloney leukemia virus reported 23 years ago [19] on which the construct had been based. In those structures, only the

second polar belt was present, and as the C-helix was absent, there was no 6HB in the trimer.

In order to fully understand the role of the polar residues observed stabilizing the 6HB, it is important to also know the structure of the spring-loaded, $(SU/TM^E)_3$ pre-fusion form. No such structure is available for any gamma-type retroviral Env protein, and the available structure of the beta-type HIV-1 Env protein is too distant to be a useful guide. In contrast, the structure of the fusion glycoprotein (GP) of the EBOV and MARV filoviruses in the pre-fusion $(GP1/GP2^E)_3$ spring-loaded trimer is available; in the filovirus GP, GP1 and GP2 are equivalent to the retroviral SU and TM subunits, respectively. The high structural similarity of the post-fusion $GP2^E$ 6HB suggests that the region spanning the N-helix in the gamma-type retroviral Env proteins can potentially undergo a similar change in going from the pre-fusion to the post-fusion conformation. Our helix-breaking proline scanning mutagenesis studies, targeting the *a* and *d* positions of the HR downstream of the stutter, identified that proline residues can be introduced at turns 6 and 7 of the N-helix (mutants L365P, L368P, L372P; Fig. 5A) without interfering with full-length syn1 being transported to the surface of the transfected cells. This observation strongly indicates that the mutants were well-folded and consequently had passed the folding quality control system of the cell, which was not the case for the proline mutants engineered further upstream (Y347P, L350P, L354P, M358P at turns 1-3 i.e. between the N-terminus and the stutter) or downstream (L375P and L379P at turns 8 and 9, closer to belt 2), which led to decreased transport to the cell surface (Fig. S3).

The similarities between the syn1-TM^E and filoviral ectodomain $GP2^E$ post-fusion 6HBs are displayed in Fig. 7B, where the region of the N-helix that tolerates proline mutations is highlighted, as well as the region that maps in between shorter helices (i.e., to a helix break") in the pre-fusion form of the filovirus GP2. Remarkably, this region is also located two turns downstream of the stutter. The GP1 moiety of the

filovirus GP is, however, most likely unrelated to the gamma-retrovirus SU, as the two fusion proteins have very different properties. For instance, filoviruses enter cells by fusion in late endosomes, and the EBOV GP requires further proteolytic cleavage of GP1 by cellular cathepsins after uptake into a target cell, partially removing the GP1 crown [57]. The remainder of GP1 then interacts with a second, internal receptor upon this cleavage, the Niemann-Pick C1 protein [58], and fusion ensues only at this late stage membrane. As a result, the pre-fusion form of the filovirus $(GP1/GP2^{E})_3$ spike is particularly robust and no additional stabilization was necessary for the structural studies. The removal of the mucin-like domain, which is heavily glycosylated, and complexation with neutralizing antibody fragments served to improve crystal packing, but were not required for stabilization of the pre-fusion conformation per se [43, 44]. In contrast, the syncytins act directly at the plasma membrane, and the fusion trigger is the interaction with their cell-surface receptor (reviewed in [59]). In this sense, they are more like the beta-type HIV-1 Env protein, which similarly drives fusion of the viral envelope with the plasma membrane for entry. The syncytins are also reminiscent of the paramyxoviruses and pneumoviruses, which also induce fusion of the viral envelope with the plasma membranes. These viral fusion proteins all required mutagenesis to maintain them in the pre-fusion form in order to maintain conformational epitopes to use as immunogens and to carry out structural studies [49, 50]. An initial structure of the pre-fusion form, however, guided those trials. With the syncytins we have worked backwards, devising trials by explicit homology of the post-fusion form without any information of the pre-fusion form. Our results nonetheless now point to new ways of engineering the pre-fusion form of gamma-type retroviral Env proteins, using mutants with helix-breaking residues in the identified region of the N-helix to destabilize the post-fusion form. Such a strategy may also require further mutations directly stabilizing the pre-fusion state, in addition to those in the region identified here. Possible combinations include the SU/TM $CXXC/CX_6CC$ motif to lock the inter-subunit disulfide

bond and avoid shedding of the SU subunit, and also the addition of a trimerization motif to keep the pre-fusion trimer together for a high resolution structural characterization.

4. MATERIALS AND METHODS

4.1. Expression construct design, protein expression and purification

Synthetic genes encoding the ectodomain portion of the TM subunit (TM^E) of syncytin-1 (residues 342 to 435, GenBank accession number **Q9UQF0**) and of syncytin-2 (residues 375 to 468, GenBank accession number **P60508**) were codon optimized for protein expression in bacteria and purchased from GeneArt. Cloning into the expression vector pET28-a vector (Novagen) was done using Gibson assembly Master Mix (New England Biolabs). The expression constructs contained the N-terminal six-histidine tag followed by a tobacco etch virus (TEV) cleavage site ENLYFQS (Fig. 1A).

The expression plasmids, syn1-TM^E and syn2-TM^E, were transformed into *Escherichia coli* BL21 DE3 cells (New England Biolabs), which were used as the expression host. Cell cultures were grown in Miller lysogeny broth (LB) containing 50 mg/L kanamycin (Sigma) at 37°C until the optical density at 600 nm reached 0.5. Protein expression was induced by addition of 0.5 mM of isopropyl-β-D-thiogalactopyranoside (IPTG), and three hours later the cells from 1 L of culture were collected by centrifugation at 4,000 x g for 20 minutes (Beckman Coulter centrifuge, JLA 8.1 rotor). Cells were lysed by sonication (Bioblock Scientific Vibra Cell sonicator 75041) in 50 ml of buffer containing 50mM Tris pH 8, 150mM NaCl, 1mM ethylenediaminetetraacetic acid (EDTA), and 0.2mM phenylmethylsulfonyl fluoride (PMSF). The insoluble fraction was removed by two rounds of centrifugation (3,000 x g for 20 minutes (Eppendorf 5810R centrifuge), and 50,000 x g for 30 minutes (Beckman Coulter centrifuge, JA25.50 rotor)). The clarified lysate was filtered through a 0.22µm filter (Millipore) before being loaded to 1ml HisTrap Excel column (GE Healthcare). Protein purification was performed according to the manufacturer's protocol. Eluted material was concentrated and further purified by SEC on a Superdex 75 16/60 column

(GE Healthcare) using 10mM Tris pH 8, 100mM NaCl as running buffer. Protein concentration was calculated using the extinction coefficient (calculated on ExPASy server [60]) of 0.6 ml/mg and 0.7 ml/mg for syn1-TM^{E} and syn2-TM^{E}, respectively.

4.2. Protein crystallization

Purified syn1-TM^{E} protein was concentrated to 10 mg/ml in Vivaspin PES cutoff 10,000 kDa, and crystallized at 18°C by the sitting-drop method with a reservoir solution containing 0.1M HEPES pH 7.5, 30% v/v 2-propanol and 0.2M $MgCl_2$. Syn2-TM^{E} was concentrated to 6 mg/ml and crystallized under the same conditions in 0.1M Tris pH 8.5, 25% v/v tertiary-butanol. The crystals of both proteins were harvested in their mother liquor solutions containing 25% glycerol as a cryoprotectant, and flash-frozen in liquid nitrogen.

4.3. X-ray diffraction data collection and structure determination

X-ray diffraction data were collected at the Proxima 2 beamline at the French national synchrotron facility SOLEIL (Gif-sur-Yvette, France). The data were processed using XDS package [61], and the initial phases were obtained by molecular replacement with Phaser [62] in CCP4 suite [63], using the coordinates of the previously crystallized fragment of the syn2-TM [29] as a search model (PDB accession number **1Y4M**). Model refinement was done in Buster [64], and structure building in Coot [65].

4.4. Structural analyses

Coiled-coils were analyzed by program TWISTER to determine the presence and position of stutters (https://pharm.kuleuven.be/apps/biocryst/twister.php). The sequence alignments were generated in Clustal Omega [66], and the one shown in Fig. S2A was further manually modified, in the T-loop and C-helix regions, to correspond to the structural superpositions of the proteins. ESPript [67] was used to calculate residue

conservation and generate all the images showing sequence alignments. PyMol was used to create all the structure images and to perform the structural superpositions (pair_fit command) [68]. The electrostatic potentials shown on Fig. 4A were calculated with the APBS tool in PyMol [69], on a scale -5 to +5, with red depicting negative and blue positive surface potentials.

The abbreviations and the accession codes for proteins used throughout the text and in the figures are: syncytin-1 (syn1, Q9UQF0), syncytin-2 (syn2, P60508), XMRV (xenotropic murine leukemia virus-like virus, D0UFA8), MoMLV (Moloney murine leukemia virus, P03385), MPMV (Mason-Pfizer monkey virus, P07575), BLV (bovine leukemia virus, AAF97918), HTLV-1 (human T-cell leukemia virus, BBD74589.1), EBOV (Ebola virus, O11457), MARV (Marburg virus, Q6UY66), CASV (California Academy of Science virus, YP_006590086.1). These sequences in some cases do not entirely match the sequences deposited with the PDB files, because mutagenesis was performed in several cases to facilitate protein expression and purification.

Abbreviations and the accession numbers of protein sequences used in the alignment in Fig. S1 are: MPMV (Mason-Pfizer monkey virus, P07575), SRV-1 (simian retrovirus-1, P04027.1), SRV-2 (simian retrovirus-2, AAA47563.1), SMRV (squirrel monkey retrovirus, NP_041262.1), REV (reticuloendotheliosis virus, ACT75575.1), FeLV (feline leukemia virus, AAC31802.1), GA-FeSV (Gardner-Arnstein feline sarcoma virus, P03391.1), MLV (murine leukemia virus, AAA46526.1), Mo-MLV (Moloney murine leukemia virus, P03385), XMRV (xenotropic murine leukemia virus, AEI59730.1), GALV (gibbon ape leukemia virus, ALV83310.1), WMSV (Woolly monkey sarcoma virus, ALV83313.1), KoRV (koala retrovirus, YP_008169844.1), BLV (bovine leukemia virus, AAF97918.1), HTLV-1 (human T-lymphotropic virus 1, BBA30582.1), STLV-1 (simian T-lymphotropic virus 1, Q03817.1), HTLV-2 (human T-lymphotropic virus 2, AAD34843.1), STLV-2 (simian T-lymphotropic virus 2, NP_056912.1), HTLV-3 (human T-lymphotropic virus 3, ACF40914.1), STLV-3 (simian T-lymphotropic virus 3

CAA61321.1), ALV (avian leukosis virus, ANF29601.1), AMV (avian myeloblastosis virus, AAA46303.1), RSV (Rous sarcoma virus, AAC08989.1), JSRV (Jaagsiekte sheep retrovirus, AAK38688.1), MMTV (mouse mammary tumor virus, AAF31475.1), BIV (bovine immunodeficiency virus, NP_040566.1), JDV (Jembrana disease virus, AAA64390.1), CAEV (caprine arthritis encephalitis virus, AAX81904.1), VMV (Visna/maedi virus, AAB25463.1), FIV (feline immunodeficiency virus, CAA40321.1), HIV-1 (human immunodeficiency virus 1, AAL01570.1), HIV-2 (human immunodeficiency virus 2, AAA76847.1), SIV (simian immunodeficiency virus, CDZ86805.1), and PLV-14 (puma lentivirus 14, AAA67170.1).

4.5. Temperature melts of syn1- and syn2-TM^{E}

Syn1- and syn2-TM^{E} were diluted to a concentration of 20μM in 100mM sodium-phosphate buffer pH 7 containing 2M guanidinium-chloride. The intrinsic fluorescence signal was measured as a function of increasing temperature in Prometheus NT.48 fluorimeter (Nanotemper), with 80% excitation light intensity and 2C°/minute temperature ramp.

Intrinsic fluorescence signal expressed by the 350nm/330nm emission ratio, which increases as the proteins unfold, is plotted as a function of temperature on Figure 4B. The plots are one of the three independent data collections that were performed for each protein.

4.6. Functional studies: syn1 mediated cell:cell fusion

4.6.1. Plasmids and cells

Plasmids pQCXIP-GFP1-10 (Addgene plasmid #68715) and pQCXIP-BSR-GFP11 (Addgene plasmid #68716), encoding a GFP variant residues, 1 to 155 (strands β1-10) and residues 156 to 172 (strand β11), respectively, were gifts from Yutaka Hata, and are described in [51]. The synthetic gene encoding the full-length syn1 (residues 1 to 538,

GenBank accession number **Q9UQF0**) was codon optimized for protein expression in mammalian cells, cloned into pcDNA3.1 expression vector, and purchased as such from GenScript. The expression construct containing the affinity tag, residues D **WSHPQFEK** *GGSA* **WSHPQFEK,** at the syn1 N-terminus was created using standard molecular biology protocols, and is referred to as the STsyncytin-1 (the two strep tags (ST) are highlighted in bold, while the short linker separating them is in italics). The single-point mutations were introduced into this construct by the 'QuikChange II Site-directed Mutagenesis Kit' from Agilent Technologies.

HEK293T platinum-E retroviral packaging cell line ('Plat-E' from Cell Biolabs), HEK293T cells (293T, ATCC) and derivatives were cultured in DMEM medium supplemented with Glutamax (Gibco), 10% fetal bovine serum (FBS) and 1% Penicillin-Streptomycin (PenStrep). Puromycin resistant cells were maintained in the presence of 1 μg/ml puromycin. The ExpiCHO cells (Thermo Fisher®) were grown in in ExpiCHO media at 37°C with 8% CO_2, in 50ml tubes or flasks, on the orbital shaker InforsHT Celltron at 130 rotations per minute (rpm).

4.6.2. Cell-surface expression of STsyncytin-1 variants analyzed by FACS

The cell-surface expression of the syncytin-1 variants was measured by FACS analysis of non-permeabilized ExpiCHO cells transfected with the same expression constructs as used for the fusion assays. These vectors were engineered to have the affinity strep-tag at the N-terminus of the protein, thus being accessible for binding to the antibodies used in FACS and immunofluorescence experiments. 293T cells could not be used for FACS analysis because they fused forming syncytia, complicating cell sorting and counting.

ExpiCHO cells were transiently transfected with the STsyncytin-1 plasmids, according to the manufacturer's protocol and using ExpifectamineCHO transfection reagent and 2 μg of plasmid DNA per 2ml of cells diluted to density of $3x10^6$ cells/ml.

Cells were grown in 50 ml conical tubes (Corning), and shaken in a rack at 37°C with 8% CO_2 at 130 rpm. Six hours post-transfection, the cells were centrifuged for 3 minutes at 200 x g, washed three times with ExpiCHO media and left to grow overnight in 2ml of fresh ExpiCHO media. Eighteen hours later the cells were counted, pelleted by centrifugation at 200 x g for 3 minutes, and treated with 1ml of 4% paraformaldehyde diluted in PBS for 15 minutes at 4°C. The rest of the staining procedure was performed in round-bottom 96-well plates; 100,000 cells were plated per well and in duplicate for each sample. After washing the cells with PBS containing 1% BSA (the washing buffer (WB)), the primary, StrepMAB-Immo antibody (IBA Biosciences, #2-1517-001) that recognizes the *SA*WSHPQFEK sequence in the strep tag at the N-terminus of syncytin-1, was incubated for 30 minutes at room temperature (RT). The antibody was diluted 1:5,000 in WB, and 0.05% saponin for the samples that required permeabilization (Fig. S3). After 2 washes with WB, the samples were incubated with a goat anti-Mouse IgG (H+L) highly cross-adsorbed secondary antibody conjugated to Alexa Fluor 647 (Invitrogen) that was diluted 1:500 in WB, for 30 minutes at RT, followed by two washes in WB. FACS acquisition on 10,000 cells per sample was performed with the Attune NxT Flow Cytometer (Thermo) for Y347P, L350P, L354P and M358P variants (Fig. S3A), and on BD FACSCanto™ cytometer (BD Biosciences) for the remaining samples (Fig. 5, Fig. S3B, C). Two different cytometers had to be used due to technical problems and the unavailability of the latter instrument for experiments involving Y347P, L350P, L354P syncytin-1 variants. All the data were analyzed via FlowJo software.

4.6.3. Cell-cell fusion assay (293T cells)

For murine retroviral vector production, the Plat-E retroviral packaging cell line was transfected with pQCXIP-BSR-GFP11 or pQCXIP-GFP1-10 using calcium chloride at 70% confluency. Media was changed 12h after transfection. Retroviral vectors were

harvested at three time points (36h, 48h and 72h after transfection) and ultracentrifuged at 22,000 x g for 1h at 4°C. For lentiviral transduction, $2x10^4$ HEK293T cells were resuspended in 150 µl of DMEM-HEPES and 15 µl of retroviral vector was added. Cells were agitated 30 seconds every 5 minutes for 2h at 37°C in a Thermomixer, and then transferred to 6-well plates with complete DMEM, and incubated for 2 days at 37°C prior to addition of puromycin. The two cell-lines stably expressing the GFP parts are referred to as 293T-GFP^{1-10} and 293T-GFP^{11}.

Transient transfection of the 293T-GFP^{1-10} and 293T-GFP^{11} cell lines with the STsyncytin-1 constructs was performed using Lipofectamine® 2000 (Invitrogen) following the manufacturer's protocol. Briefly, after a PBS wash, the 293T-GFP^{1-10} and 293T-GFP^{11} cells were detached with PBS-EDTA (Sigma), and resuspended in DMEM medium supplemented with Glutamax, 10% FBS and 1µg/ml puromycin. The two cell types were mixed in 1:1 ratio at a final density of $6x10^5$ cells/ml, and 100µl of such cell mixture was seeded in a 96-well cell-culture flat-bottom plate ($6x10^4$ cells/well). The same amount of cells was seeded in Lab-Tek chamber slides (#177402, ThermoFisher) for immunofluorescence analysis of syncytin surface expression (section 4.6.4). A transfection mixture containing 50µl of OptiMEM (Gibco), 100ng of STsyncytin-1 plasmid DNA and 0.3µl of the lipofectamine reagent was prepared for each sample, per well. 50µl of the mixture was added, and the cells were allowed to fuse for next 24 hours when the images were collected on live cells using the 'EVOS FL Cell Imaging System' set up to record GFP fluorescence. An ImageJ script created by J.B. for the ImageJ software [70], was applied to 1) convert the images collected on the GFP channel from the *tif format to black and white images, green signal being replaced by black color (Fig. 6), and 2) to integrate the surface occupied by the black fields and represent it as the percentage of total area, thereby giving what is referred to as 'the fusion index'. Two fields (1,142.84 x 857.13 µm) per experiment were taken for quantifications. The experiment was repeated twice, and the bar graph shown in Fig. 7A is an average of 4 measurements.

The fusion was calculated as follows: the raw data obtained by recording GFP fluorescence (Fig. 6, left column) were processed so that the green signal (fusion events) was replaced with black color, and the background signal was colored in white, as shown in the middle column. The 'fusion index' was calculated as the percent of black area in the total visualized surface. The right column on Fig. 6 represents the same cell population seeded on a separate slide for immunofluorescence staining of the surface-expressed syncytin variants (red staining).

4.6.4. Surface-expression of STsyncytin-1 variants analyzed by immunofluorescence

The 293T-GFP$^{1-10}$ and 293T-GFP^{11} cells were seeded in Lab-Tek chamber slides and transfected with STsyncytin-1 constructs as described in section 4.6.3. Twenty-four hours later the cells were fixed in 4% paraformaldehyde in PBS for 15 minutes. Free aldehydes were blocked by incubation with 40mM NH_4Cl in PBS for 5 minutes. Non-specific binding was blocked with 1% FBS in PBS containing 0.02% sodium azide for 10 minutes, followed by 1 hour, room temperature incubation with the StrepMAB-Immo antibody (IBA Lifesciences, Cat. No. 2-1517-001) diluted 3,000 times in blocking buffer. After 3 washes in PBS, the secondary antibody, goat antibody against the mouse IgG conjugated with Alexa Fluor 555 (Invitrogen, #A28180), diluted 1:2,000 was added and incubated for 1 hour. Nuclei were stained with Hoechst stain (Invitrogen) at 1µg/mL in PBS for 5 minutes before mounting the coverslips onto 76 x 26mm slides (Thermofisher) using Fluoromount-G mounting media (Southern Biotech). Images were taken with a Leica DM LB microscope and analyzed with ImageJ software.

4.6.5. Statistical analyses

In Fig. 7A, surface expression results of STsyncytin-1 variants are represented as the mean of three independent experiments, fusion indices as the mean of four measurements from two independent experiments, in all cases normalized to the value

of the positive control (syn1-WT) ± SEM. One-way ANOVA and Dunnett's multiple comparisons tests were used to determine the statistical significance of each result in comparison with syn1-WT surface expression or fusion index (*$p<0.05$; **$p<0.01$; ****$p<0.0001$). A p value of <0.05 was accepted as indicative of a statistically significant difference. The fusion index bars are missing from most of the samples because syncytia were not observed. The ****$p<0.0001$ values were calculated for the fusion indexes for all the samples and are not plotted on the graph for clarity reasons.

ACCESSION NUMBERS: Coordinates and structure factors have been deposited in the Protein Data Bank with the accession numbers **6RX1** and **6RX3**, for syn1- and syn2-TM^{E}, respectively.

5. ACKNOWLEDGMENTS

This work was funded by the European Research council (ERC 'CelCelFus' grant #340371) awarded to F.A.R, and recurrent funding from the Institut Pasteur (IP) and CNRS. We thank Ahmed Haouz and staff at the IP core facility for protein crystallization, and staff at the Proxima-2 beamline at the SOLEIL synchrotron for assistance with data collection and processing. We also thank Ignacio Fernandez for helpful discussions and critical reading of the manuscript, Delphine Brun for technical assistance, and Etienne Simon-Loriere for discussions and help with phylogenetic analyses.

6. REFERENCES

[1] Lavialle C, Cornelis G, Dupressoir A, Esnault C, Heidmann O, Vernochet C, et al. Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation. Philos Trans R Soc Lond B Biol Sci. 2013;368:20120507.
[2] Dupressoir A, Lavialle C, Heidmann T. From ancestral infectious retroviruses to bona fide cellular genes: role of the captured syncytins in placentation. Placenta. 2012;33:663-71.
[3] Soe K, Andersen TL, Hobolt-Pedersen AS, Bjerregaard B, Larsson LI, Delaisse JM. Involvement of human endogenous retroviral syncytin-1 in human osteoclast fusion. Bone. 2011;48:837-46.
[4] Dewannieux M, Heidmann T. Endogenous retroviruses: acquisition, amplification and taming of genome invaders. Curr Opin Virol. 2013;3:646-56.
[5] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409:860-921.
[6] Cornelis G, Souquere S, Vernochet C, Heidmann T, Pierron G. Functional conservation of the lncRNA NEAT1 in the ancestrally diverged marsupial lineage: Evidence for NEAT1 expression and associated paraspeckle assembly during late gestation in the opossum Monodelphis domestica. RNA Biol. 2016;13:826-36.
[7] Cornelis G, Funk M, Vernochet C, Leal F, Tarazona OA, Meurice G, et al. An endogenous retroviral envelope syncytin and its cognate receptor identified in the viviparous placental Mabuya lizard. Proc Natl Acad Sci U S A. 2017;114:E10991-E1000.
[8] Vargiu L, Rodriguez-Tome P, Sperber GO, Cadeddu M, Grandi N, Blikstad V, et al. Classification and characterization of human endogenous retroviruses; mosaic forms are common. Retrovirology. 2016;13:7.
[9] Jern P, Sperber GO, Blomberg J. Use of endogenous retroviral sequences (ERVs) and structural markers for retroviral phylogenetic inference and taxonomy. Retrovirology. 2005;2:50.
[10] Blond JL, Beseme F, Duret L, Bouton O, Bedin F, Perron H, et al. Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family. J Virol. 1999;73:1175-85.
[11] Blaise S, de Parseval N, Benit L, Heidmann T. Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. Proc Natl Acad Sci U S A. 2003;100:13013-8.
[12] Blaise S, Ruggieri A, Dewannieux M, Cosset FL, Heidmann T. Identification of an envelope protein from the FRD family of human endogenous retroviruses (HERV-FRD) conferring infectivity and functional conservation among simians. J Virol. 2004;78:1050-4.
[13] Henzy JE, Johnson WE. Pushing the endogenous envelope. Philos Trans R Soc Lond B Biol Sci. 2013;368:20120506.
[14] Rey FA, Lok SM. Common Features of Enveloped Viruses and Implications for Immunogen Design for Next-Generation Vaccines. Cell. 2018;172:1319-34.
[15] Harrison SC. Mechanism of membrane fusion by viral envelope proteins. Adv Virus Res. 2005;64:231-61.
[16] Li K, Zhang S, Kronqvist M, Wallin M, Ekstrom M, Derse D, et al. Intersubunit disulfide isomerization controls membrane fusion of human T-cell leukemia virus Env. J Virol. 2008;82:7135-43.
[17] Benhaim M, Lee KK. Single-Molecule Analysis of a Viral Fusion Protein Illuminates a Fusion-Active Intermediate State. Cell. 2018;174:775-7.
[18] Leikin S, Parsegian VA, Rau DC, Rand RP. Hydration forces. Annu Rev Phys Chem. 1993;44:369-95.
[19] Fass D, Harrison SC, Kim PS. Retrovirus envelope domain at 1.7 angstrom resolution. Nat Struct Biol. 1996;3:465-9.

[20] Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC. Atomic structure of the ectodomain from HIV-1 gp41. Nature. 1997;387:426-30.
[21] Yang ZN, Mueser TC, Kaufman J, Stahl SJ, Wingfield PT, Hyde CC. The crystal structure of the SIV gp41 ectodomain at 1.47 A resolution. J Struct Biol. 1999;126:131-44.
[22] Aydin H, Cook JD, Lee JE. Crystal structures of beta- and gammaretrovirus fusion proteins reveal a role for electrostatic stapling in viral entry. J Virol. 2014;88:143-53.
[23] Aydin H, Smrke BM, Lee JE. Structural characterization of a fusion glycoprotein from a retrovirus that undergoes a hybrid 2-step entry mechanism. FASEB J. 2013;27:5059-71.
[24] Kobe B, Center RJ, Kemp BE, Poumbourios P. Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. Proc Natl Acad Sci U S A. 1999;96:4319-24.
[25] Lamb D, Schuttelkopf AW, van Aalten DM, Brighty DW. Charge-surrounded pockets and electrostatic interactions with small ions modulate the activity of retroviral fusion proteins. PLoS Pathog. 2011;7:e1001268.
[26] Ward AB, Wilson IA. The HIV-1 envelope glycoprotein structure: nailing down a moving target. Immunol Rev. 2017;275:21-32.
[27] Pancera M, Changela A, Kwong PD. How HIV-1 entry mechanism and broadly neutralizing antibodies guide structure-based vaccine design. Curr Opin HIV AIDS. 2017;12:229-40.
[28] Sanders RW, Vesanen M, Schuelke N, Master A, Schiffner L, Kalyanaraman R, et al. Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. J Virol. 2002;76:8875-89.
[29] Renard M, Varela PF, Letzelter C, Duquerroy S, Rey FA, Heidmann T. Crystal structure of a pivotal domain of human syncytin-2, a 40 million years old endogenous retrovirus fusogenic envelope gene captured by primates. J Mol Biol. 2005;352:1029-34.
[30] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157:105-32.
[31] Sonnhammer EL, von Heijne G, Krogh A. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol. 1998;6:175-82.
[32] Holm L, Laakso LM. Dali server update. Nucleic Acids Res. 2016;44:W351-5.
[33] Sonigo P, Barker C, Hunter E, Wain-Hobson S. Nucleotide sequence of Mason-Pfizer monkey virus: an immunosuppressive D-type retrovirus. Cell. 1986;45:375-85.
[34] Maerz AL, Center RJ, Kemp BE, Kobe B, Poumbourios P. Functional implications of the human T-lymphotropic virus type 1 transmembrane glycoprotein helical hairpin structure. J Virol. 2000;74:6614-21.
[35] Oakley MG, Kim PS. A buried polar interaction can direct the relative orientation of helices in a coiled coil. Biochemistry. 1998;37:12603-10.
[36] Lumb KJ, Kim PS. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. Biochemistry. 1998;37:13042.
[37] Mangeney M, Renard M, Schlecht-Louf G, Bouallaga I, Heidmann O, Letzelter C, et al. Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins. Proc Natl Acad Sci U S A. 2007;104:20534-9.
[38] Gallaher WR. Similar structural models of the transmembrane proteins of Ebola and avian sarcoma viruses. Cell. 1996;85:477-8.
[39] Koellhoffer JF, Dai Z, Malashkevich VN, Stenglein MD, Liu Y, Toro R, et al. Structural characterization of the glycoprotein GP2 core domain from the CAS virus, a novel arenavirus-like species. J Mol Biol. 2014;426:1452-68.
[40] Cook JD, Lee JE. The secret life of viral entry glycoproteins: moonlighting in immune evasion. PLoS Pathog. 2013;9:e1003258.

[41] Weissenhorn W, Carfi A, Lee KH, Skehel JJ, Wiley DC. Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. Mol Cell. 1998;2:605-16.
[42] Koellhoffer JF, Malashkevich VN, Harrison JS, Toro R, Bhosle RC, Chandran K, et al. Crystal structure of the Marburg virus GP2 core domain in its postfusion conformation. Biochemistry. 2012;51:7665-75.
[43] Zhao Y, Ren J, Harlos K, Jones DM, Zeltina A, Bowden TA, et al. Toremifene interacts with and destabilizes the Ebola virus glycoprotein. Nature. 2016;535:169-72.
[44] Lee JE, Fusco ML, Hessell AJ, Oswald WB, Burton DR, Saphire EO. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. Nature. 2008;454:177-82.
[45] Hashiguchi T, Fusco ML, Bornholdt ZA, Lee JE, Flyak AI, Matsuoka R, et al. Structural basis for Marburg virus neutralization by a cross-reactive human antibody. Cell. 2015;160:904-12.
[46] Brown JH, Cohen C, Parry DA. Heptad breaks in alpha-helical coiled coils: stutters and stammers. Proteins. 1996;26:134-45.
[47] Strelkov SV, Burkhard P. Analysis of alpha-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. J Struct Biol. 2002;137:54-64.
[48] Pallesen J, Wang N, Corbett KS, Wrapp D, Kirchdoerfer RN, Turner HL, et al. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. Proc Natl Acad Sci U S A. 2017;114:E7348-E57.
[49] Krarup A, Truan D, Furmanova-Hollenstein P, Bogaert L, Bouchier P, Bisschop IJ, et al. A highly stable prefusion RSV F vaccine derived from structural analysis of the fusion mechanism. Nat Commun. 2015;6:8143.
[50] Battles MB, Mas V, Olmedillas E, Cano O, Vazquez M, Rodriguez L, et al. Structure and immunogenicity of pre-fusion-stabilized human metapneumovirus F glycoprotein. Nat Commun. 2017;8:1528.
[51] Kodaka M, Yang Z, Nakagawa K, Maruyama J, Xu X, Sarkar A, et al. A new cell-based assay to evaluate myogenesis in mouse myoblast C2C12 cells. Exp Cell Res. 2015;336:171-81.
[52] Buchrieser J, Degrelle SA, Couderc T, Nevers Q, Disson O, Manet C, et al. IFITM proteins inhibit placental syncytiotrophoblast formation and promote fetal demise. Science. 2019;365:176-80.
[53] Blond JL, Lavillette D, Cheynet V, Bouton O, Oriol G, Chapel-Fernandes S, et al. An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. J Virol. 2000;74:3321-9.
[54] Chen CP, Chen LF, Yang SR, Chen CY, Ko CC, Chang GD, et al. Functional characterization of the human placental fusogenic membrane protein syncytin 2. Biol Reprod. 2008;79:815-23.
[55] Cheynet V, Ruggieri A, Oriol G, Blond JL, Boson B, Vachot L, et al. Synthesis, assembly, and processing of the Env ERVWE1/syncytin human endogenous retroviral envelope. J Virol. 2005;79:5585-93.
[56] Thomas A, Roth MJ. Analysis of cysteine mutations on the transmembrane protein of Moloney murine leukemia virus. Virology. 1995;211:285-9.
[57] Chandran K, Sullivan NJ, Felbor U, Whelan SP, Cunningham JM. Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. Science. 2005;308:1643-5.
[58] Carette JE, Raaben M, Wong AC, Herbert AS, Obernosterer G, Mulherkar N, et al. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. Nature. 2011;477:340-3.

[59] Greenwood AD, Ishida Y, O'Brien SP, Roca AL, Eiden MV. Transmission, Evolution, and Endogenization: Lessons Learned from Recent Retroviral Invasions. Microbiol Mol Biol Rev. 2018;82.
[60] Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003;31:3784-8.
[61] Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010;66:125-32.
[62] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007;40:658-74.
[63] Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67:235-42.
[64] Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, et al. BUSTER version 2.8.0. Cambridge, United Kingdom: Global Phasing Ltd.; 2009.
[65] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66:486-501.
[66] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.
[67] Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics. 1999;15:305-8.
[68] DeLano WL. The PyMOL Molecular Graphics System. San Carlos, CA, USA: DeLano Scientific; 2002.
[69] Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. 2004;32:W665-7.
[70] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9:671-5.

FIGURE LEGENDS

Figure 1. (A) Organization of the full-length syncytin proteins. The syncytin main features are displayed as boxes on the linear diagram. Black horizontal lines at the top show the extent of the SU and TM subunits, together with the TM ectodomain (TM^E). The boundaries of the crystallized constructs and the residues resolved in the structures are marked above and below the rectangular diagram of the full-length syncytin-1 and syncytin-2, respectively. The residues limiting different regions within SU and TM are indicated with numbers on the scheme of the full-length protein. The inter-subunit disulfide bond is represented with dashed lines, and the intra-TM disulfide bond within CX_6CC motif in T-loop as an arc symbol. **(B) Sequence alignment of syncytin-1 and syncytin-2 TM subunits.** The residues are colored according to their conservation: identical residues in white font on a red background, non-conserved residues in black font, and the residues with similar physicochemical properties in red font on a white background. The boundaries of the expression constructs are marked with black vertical lines, and the regions not resolved in the structures with dashed lines. The helical register (*abcdefg*) and the turns in N- and C-helices are plotted above and below the sequence alignment. The stutter residues are underlined. The N-helix asparagines that coordinate chloride ions are labeled with 'Cl^-', and the T-loop residues conferring the ISU + fusion properties to syncytin-2 are highlighted in yellow. Orange arrowheads indicate the boundaries of the previously crystallized syn2-TM^E.

Figure 2. Post-fusion structures of syn1-TM^E and syn2-TM^E. Our X-ray structures of syn1- and syn2-TM^E are compared with the previously reported syn2 structure (top row), and the structures of closely related MPMV and XMRV TM^E (bottom row). The structural elements forming the TM^E - N-helix (pale gray), T-loop (blue), C-helix (dark gray) - are indicated, as well as the turns in the helices. The intra-chain disulfide bonds are shown

as yellow sticks. A pale-green tube inside the N-helices follows the axis of each helix, with the stutter location marked black. A black vertical line marks the 3-fold axis of the coiled-coil. A chloride ion coordinated by asparagine residues is shown as a purple sphere. The residues conferring ISU properties to syn2-TM^{E}, and the equivalent residues in other TM^{E} have their side chains displayed as orange sticks, and their names displayed on yellow background. The conserved N-glycosylation site at the beginning of C-helix is labeled as N409, N443, and N487 in syn1-, syn2-, and MPMV TM^{E}, respectively. The PDB codes are displayed below the structure names.

Figure 3. Interactions within syncytin 6HBs. (A) Inter-chain polar interactions. Side views on syn1- and syn2-TM^{E} are shown on left and right, respectively, zooming into the belt 1 and belt 2 regions in the center. Residues belonging to the neighboring protomer are labeled with an asterisk. **(B) Syn1-TM^{E} 'vertical' string of salt bridges.** Inter-chain hydrogen bonds formed between charged residues forming 8 salt bridges are shown on the left, and the single intra-chain salt bridge is indicated on the right. **(C) Syn2-TM^{E} inter- and intra-chain salt bridges.** Hydrogen bonds formed between the charged residues forming salt bridges at the N- and C-termini of the bundle are shown. **(D) Comparison of interactions within 6HBs of syncytins and gamma-type Envs.** The residues forming belt 1 and 2 are displayed on orange and green background, respectively. The charged amino acids forming a vertical string in syn1-TM^{E}, and the charged residues forming salt bridges mapping to the same region in syn2-, MPMV and XMRV TM^{E} are indicated as red, bold letters. The numbers below the sequence alignment mark the turns in the N- and C-helices of syncytins' 6HB. Residues at positions *a* and *d* within the HRs are indicated (the *a* position in the stutter is in red).

Figure 4. (A) Electrostatic potentials of syncytin 6HBs. The electrostatic potentials of syn1- and syn-2 TM^{E} coiled-coils are plotted on the molecular surface, with the C-helix

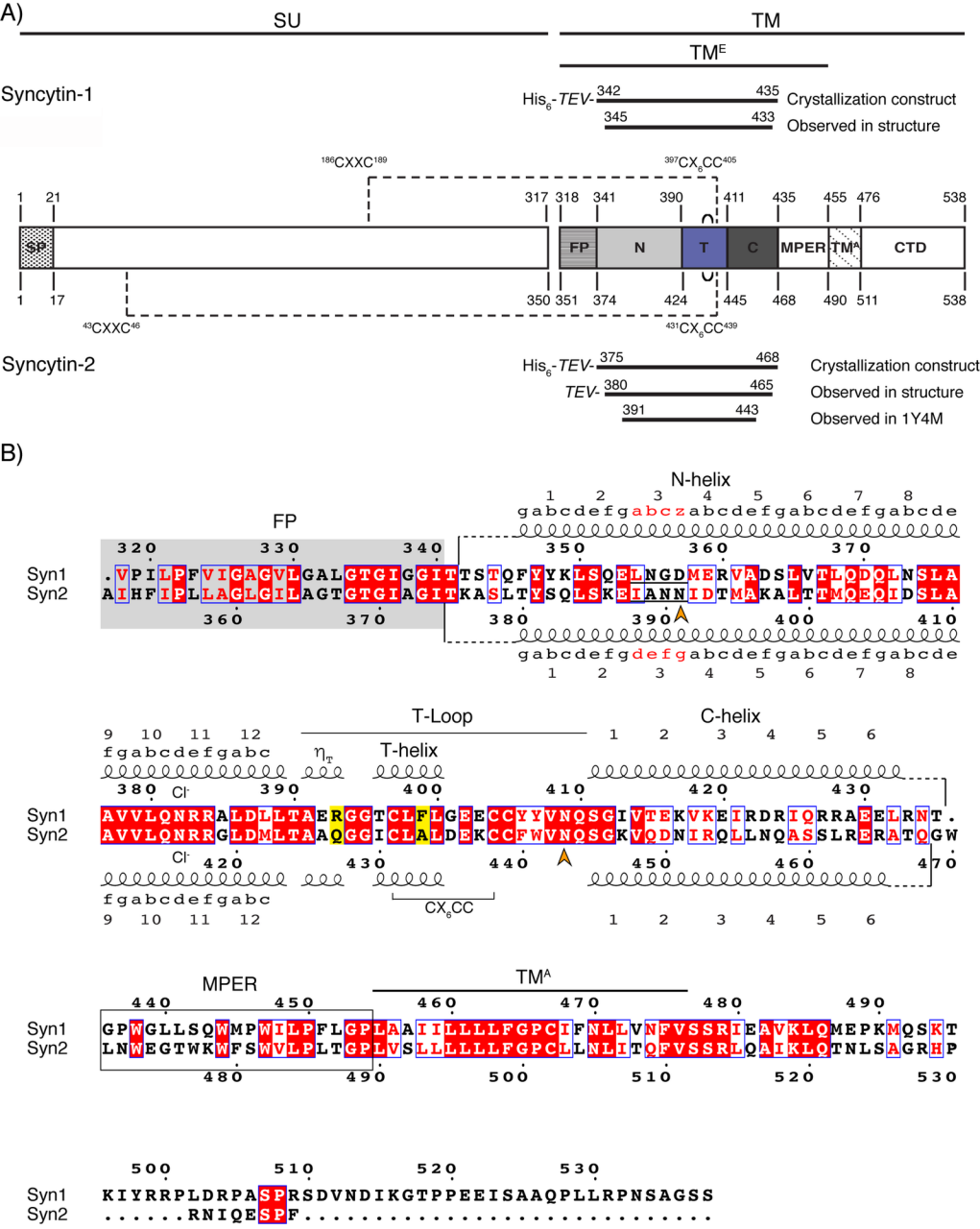
presented as cartoon, in pale grey and with numbers 1-5 designating the turns (left side of each panel). The C-helix rotated around the Y-axis for 180° is shown on the right to illustrate the complementarity of the positively charged C-helical surfaces, and the highly negatively charged groove formed by the N-helices in syn1-TM^{E}. **(B) 6HB bundle of syncytin-1 is more stable than of syncytin-2.** The graph shows unfolding of syn1- and syn2-TM^{E} displayed as increase in the intrinsic fluorescence as a function of temperature. The syn1-TM^{E} unfolds at a temperature ~20°C higher than syn2-TM^{E}.

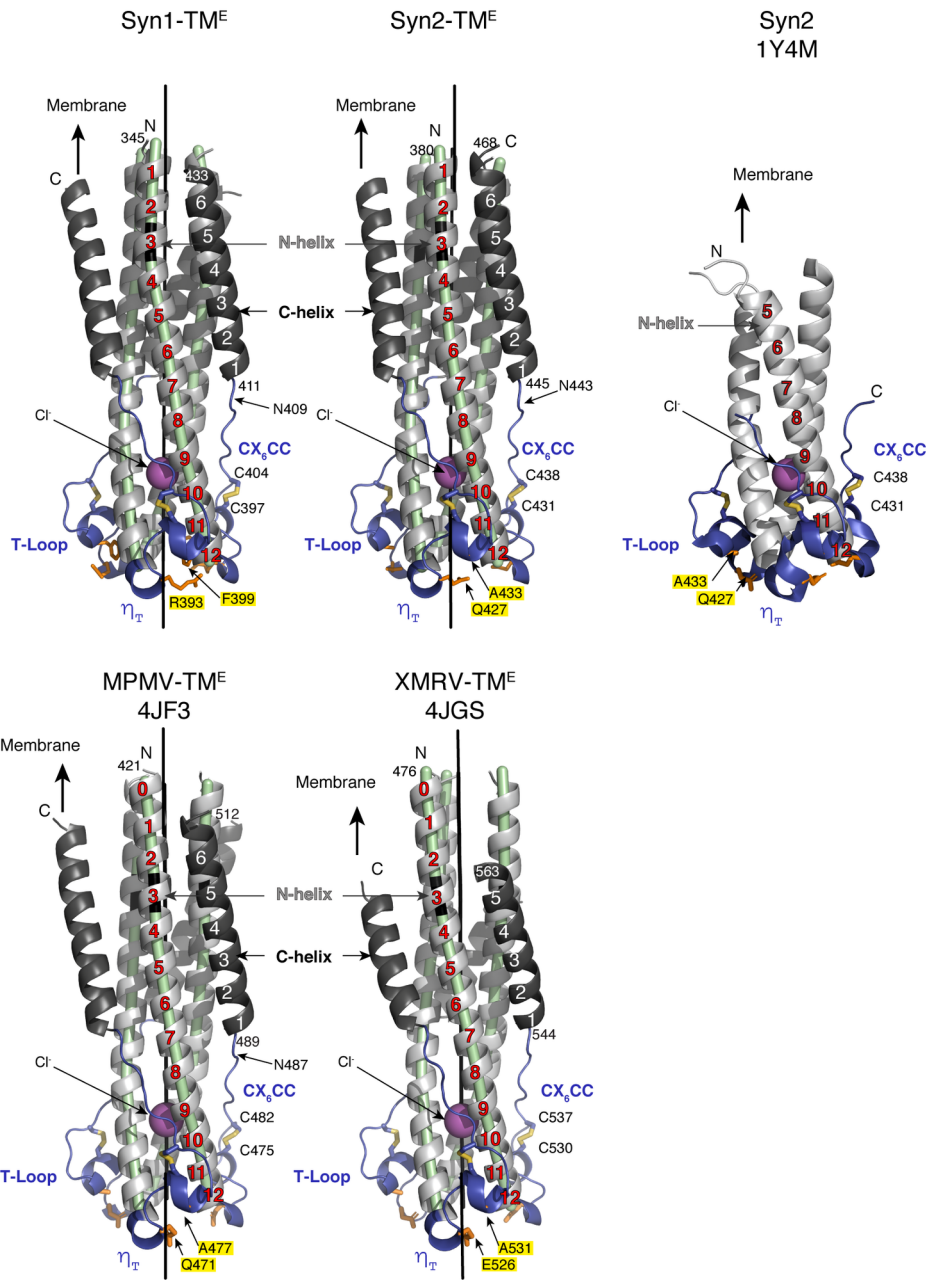
Figure 5. L365P, L368P and L372P syncytin-1 variants are expressed at the surface of ExpiCHO cells at a level similar to WT. Surface or total expression of syncytin-1 variants, marked on the left side of each row, were measured by FACS. The pink numbers in each dot-plot represent the percentage of cells that tested positive for syncytin surface or total expression. A) The proline variants L365P, L368P and L372P, were expressed at the cells surface at a level similar to the WT protein (left panel). The side-chains of L365P, L368P and L372P are shown as dark green sticks and their names are displayed in bold letters, while the other residues mutated in this study are shown as pale green sticks (right panel). B) Lower surface expression was measured for the C186S variant, which contains cysteine to serine mutation in the C^{186}XXC motif, and the variant with mutated furin site, suggesting impaired folding and / or trafficking caused by the mutation. The experimental details are given in section 4.6.2.

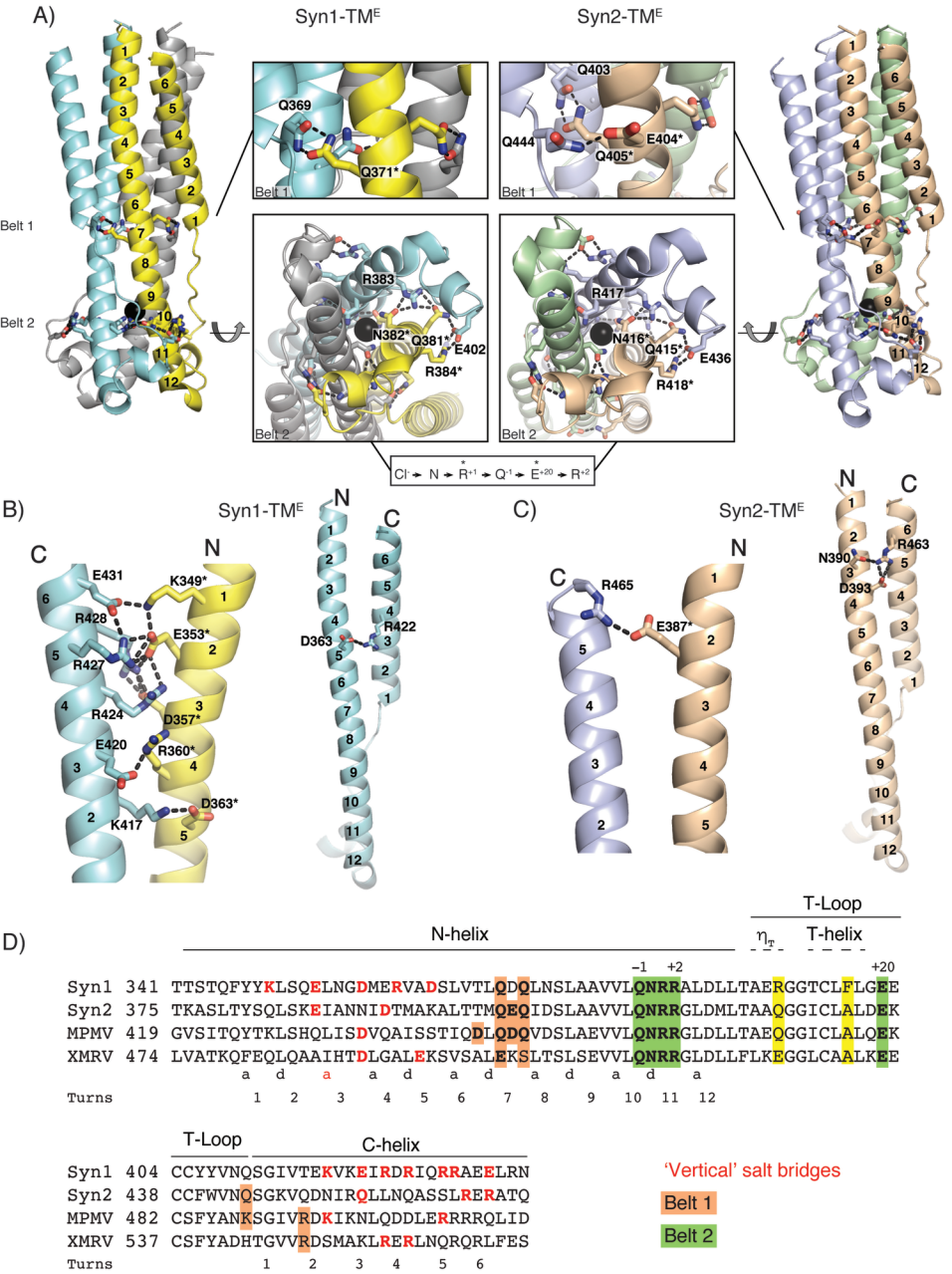
Figure 6. Syncytin-1 induced fusion of 293T cells. The fusion assay results are displayed in the first two columns (raw data and the processed data, respectively, as explained in the section 4.6.3). Results for only one of the proline variants are shown (L368P), as it is representative of all the proline variants, none of which were active in fusion (for complete data see Fig. S4). The images in the third column correspond to the syncytin-1 surface-expression (red), observed by immunofluorescence of 293T cells transfected

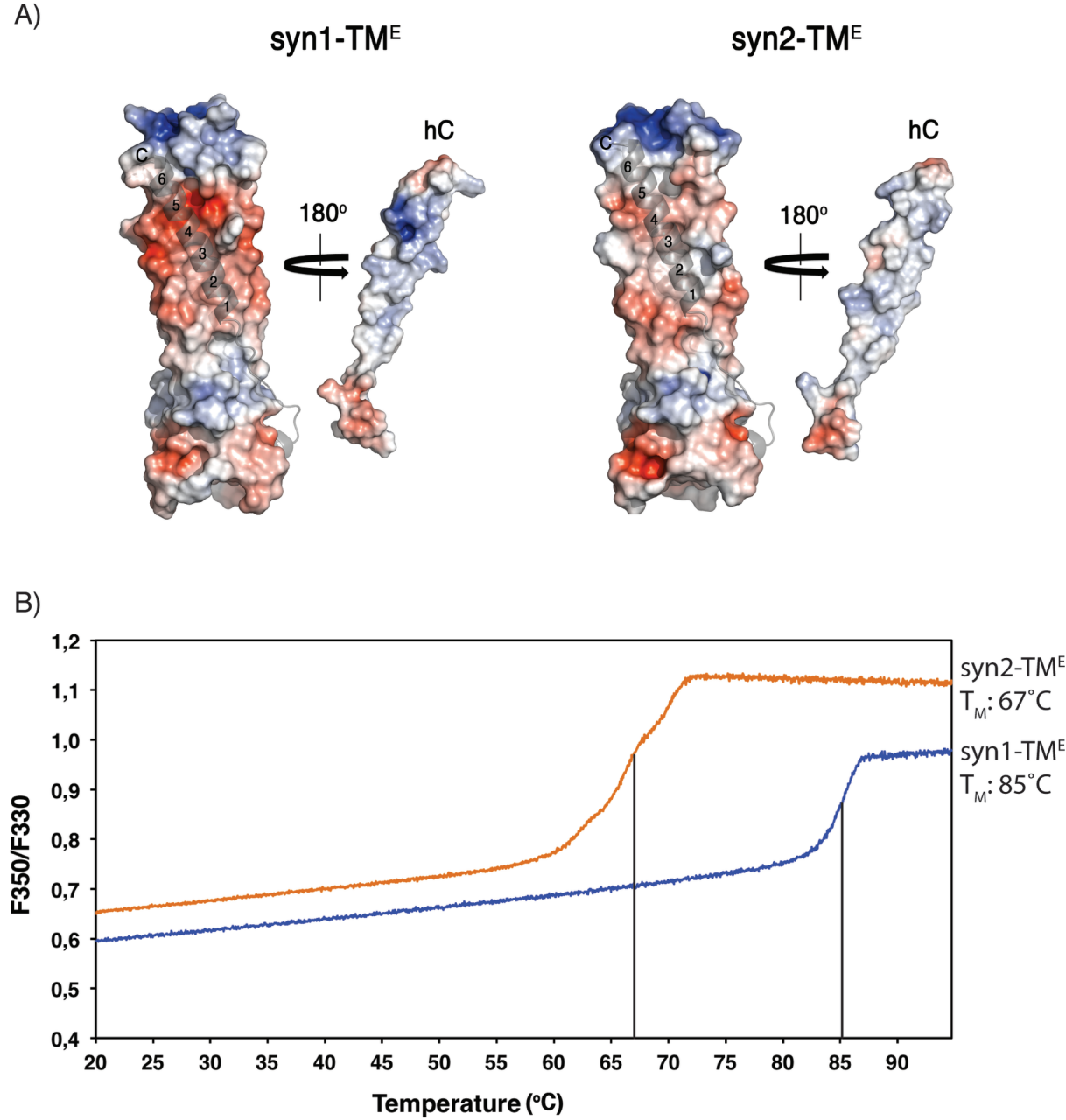
with STsyncytin-1 variants (described in section 4.6.4). The images for the fusion assay and immunofluorescence were taken by different microscopes and under different magnifications, indicated by the scale bars of 200 µm.

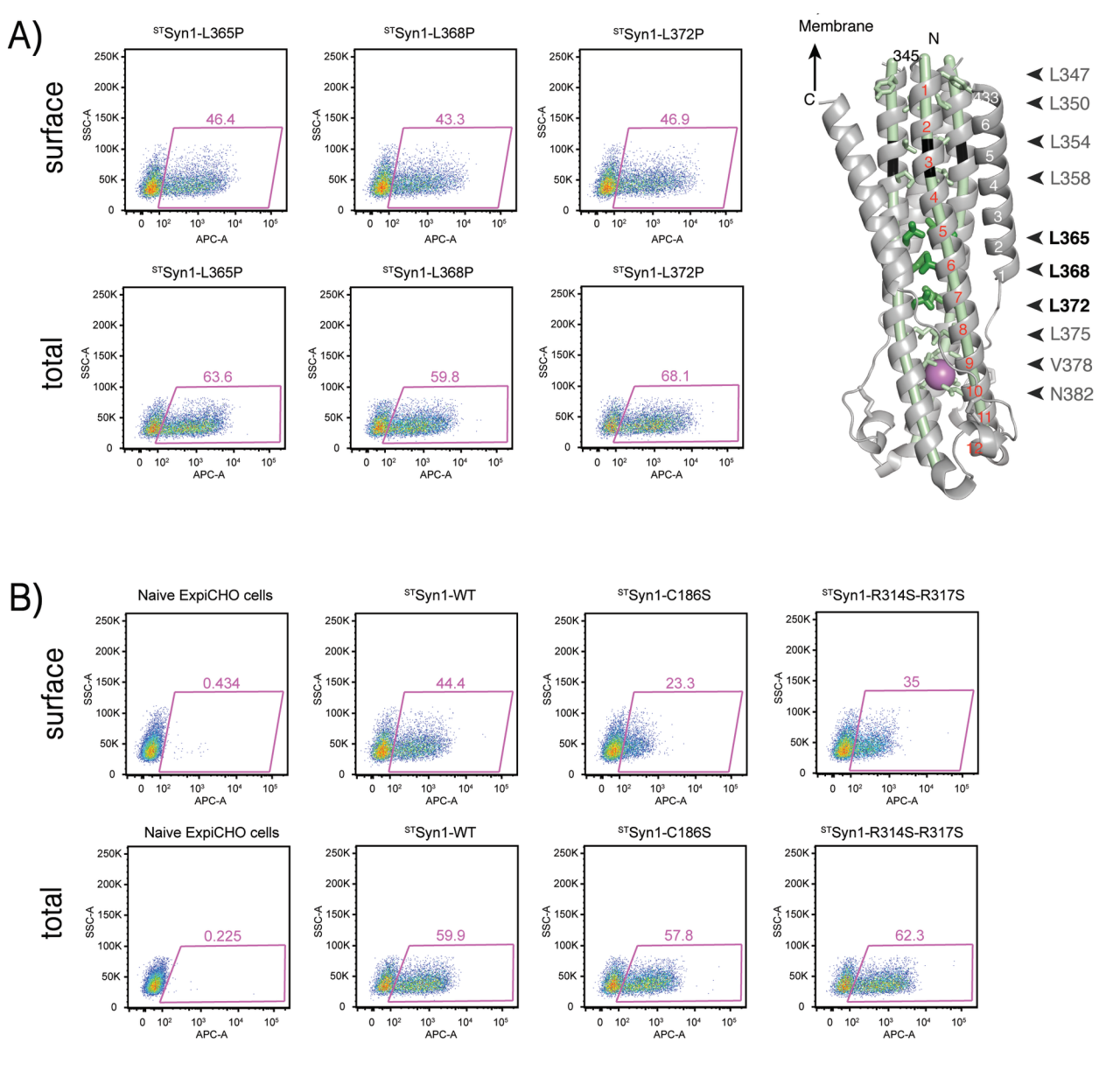
Figure 7. Prolines in fusion-inactive, correctly surface expressed syncytin-1 variants map to the same location relative to the stutter as the helix-breaks in the pre-fusion conformation of filoviral GP2. (A) The bar chart displays surface expression measured by FACS (white bars) and fusion index (dark grey bars), both set to 100% for the WT syncytin-1 and expressed relative to the WT values. The error bars indicate standard error of the mean values (SEM). Statistical analyses were done as described in section 4.6.5. **(B)** The structures are colored to highlight the stutters (black) and the regions that form helix-breaks in the pre-fusion forms of EBOV and MARV GP2 (blue). The blue region in syncytin-1 indicates the locations of the proline mutations that rendered the protein inactive, but did not impact its trafficking to the cell surface, indicating that the same region may form a helix-break in the syncytin-1 metastable state. The open brackets illustrate the conserved distance between the stutters and the predicted or already known helix-breaks in syncytin-1 and filoviral GP2, respectively. The HR2 regions are omitted for clarity.

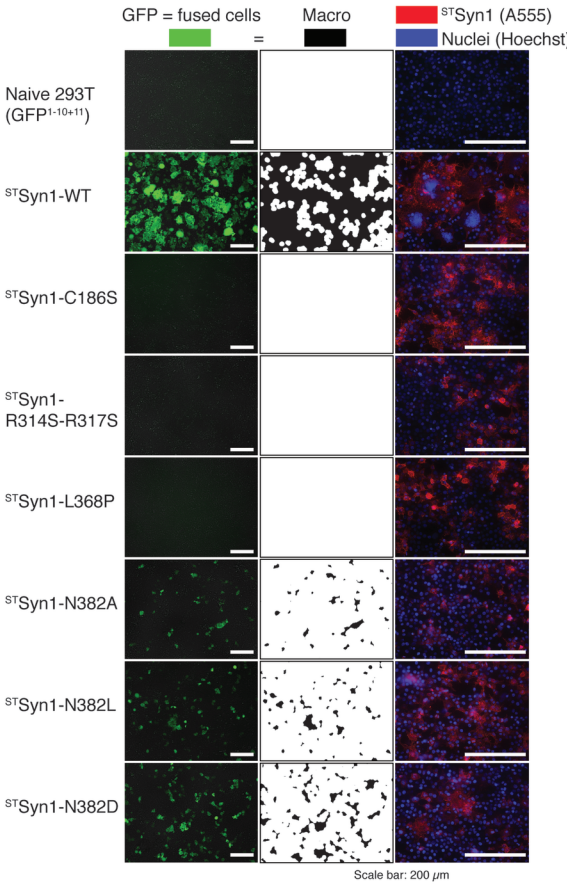

A)
SU
TM
TM^E
Syncytin-1
His6-TEV-342 435 Crystallization construct
345 433 Observed in structure
186CXXC189
397CX6CC405
SP
FP
N
T
C
MPER
TM^A
CTD
43CXXC46
431CX6CC439
Syncytin-2
His6-TEV-375 468 Crystallization construct
TEV-380 465 Observed in structure
391 443 Observed in 1Y4M
B)
N-helix
FP
Syn1
Syn2
T-Loop
C-helix
T-helix
MPER
TM^A

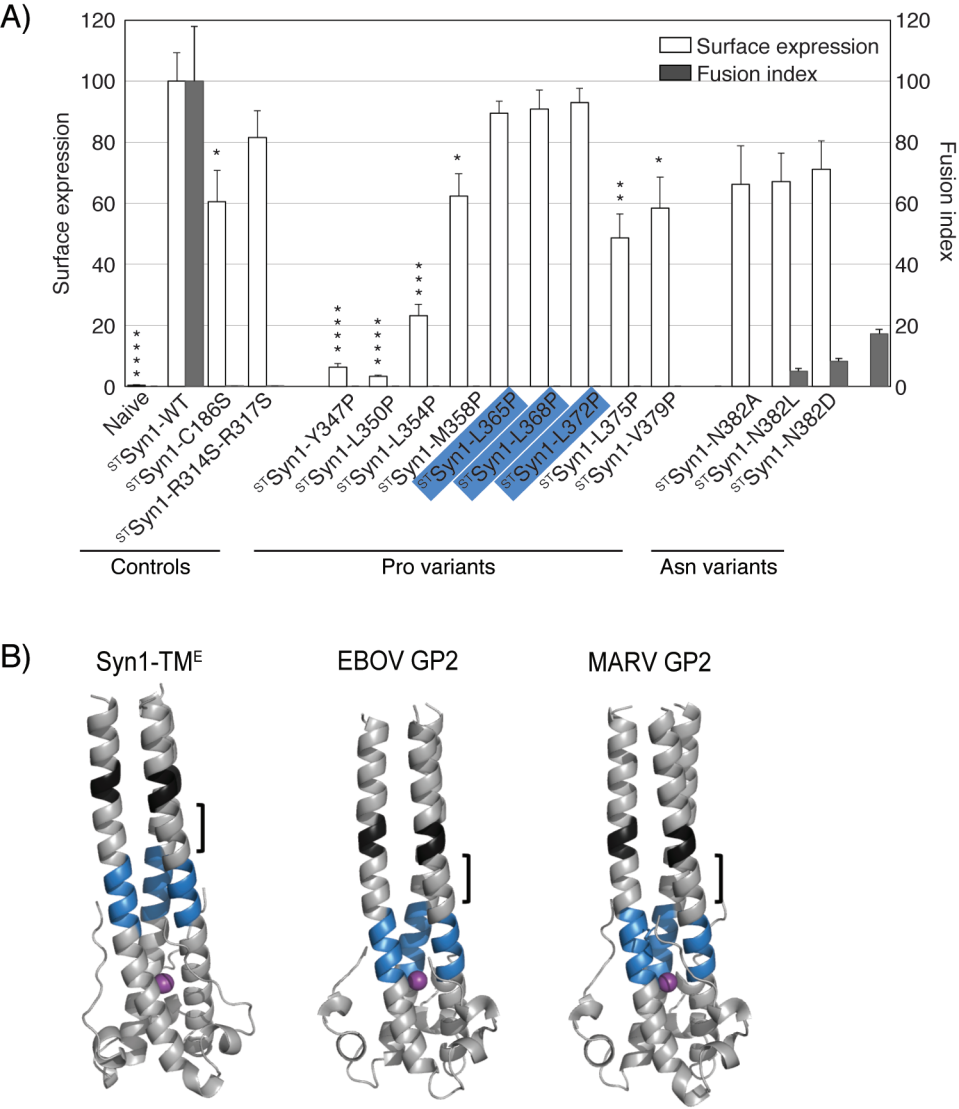

Syn1-TM^E
Syn2-TM^E
Syn2
1Y4M
Membrane
N
C
345
433
380
468
N-helix
C-helix
411
N409
445
N443
Cl-
CX6CC
C404
C438
C397
C431
T-Loop
R393
F399
A433
Q427
MPMV-TM^E
4JF3
XMRV-TM^E
4JGS
421
512
476
563
489
N487
544
C482
C475
C537
C530
A477
Q471
A531
E526


A)
Syn1-TM^E
Syn2-TM^E
Belt 1
Belt 2
Q369
Q371*
Q403
Q444
E404*
Q405*
R383
N382*
Q381*
E402
R384*
R417
N416*
Q415*
E436
R418*
Cl⁻ → N → R*⁻¹ → Q⁻¹ → E*⁻²⁰ → R⁺²
B)
Syn1-TM^E
C
N
E431
K349*
R428
E353*
R427
R424
D357*
E420
R360*
D363*
K417
D363
R422
C)
Syn2-TM^E
R465
E387*
N390
R463
D393
D)
N-helix
T-Loop
T-helix
Syn1 341 TTSTQFYYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLQNRRALDLLTAERGGTCLFLGEE
Syn2 375 TKASLTYSQLSKEIANNIDTMAKALTTMQEQIDSLAAVVLQNRRGLDMLTAAQGGICLALDEK
MPMV 419 GVSITQYTKLSHQLISDVQAISSTIQDLQDQVDSLAEVVLQNRRGLDLLTAEQGGICLALQEK
XMRV 474 LVATKQFEQLQAAIHTDLGALEKSVSALEKSLTSLSEVVLQNRRGLDLLFLKEGGLCAALKEE
a d a a d a d a d a d a
Turns 1 2 3 4 5 6 7 8 9 10 11 12
T-Loop
C-helix
Syn1 404 CCYYVNQSGIVTEKVKEIRDRIQRRAEELRN
Syn2 438 CCFWVNQSGKVQDNIRQLLNQASSLRERATQ
MPMV 482 CSFYANKSGIVRDKIKNLQDDLERRRRQLID
XMRV 537 CSFYADHTGVVRDSMAKLRERLNQRQRLFES
Turns 1 2 3 4 5 6
'Vertical' salt bridges
Belt 1
Belt 2


A)
syn1-TM^E
syn2-TM^E
hC
180º
C
6
5
4
3
2
1
B)
F350/F330
Temperature (ºC)
syn2-TM^E
T_M: 67˚C
syn1-TM^E
T_M: 85˚C


A)
surface
total
STSyn1-L365P
STSyn1-L368P
STSyn1-L372P
46.4
43.3
46.9
63.6
59.8
68.1
SSC-A
APC-A
Membrane
N
C
345
133
L347
L350
L354
L358
L365
L368
L372
L375
V378
N382
B)
Naive ExpiCHO cells
STSyn1-WT
STSyn1-C186S
STSyn1-R314S-R317S
0.434
44.4
23.3
35
0.225
59.9
57.8
62.3

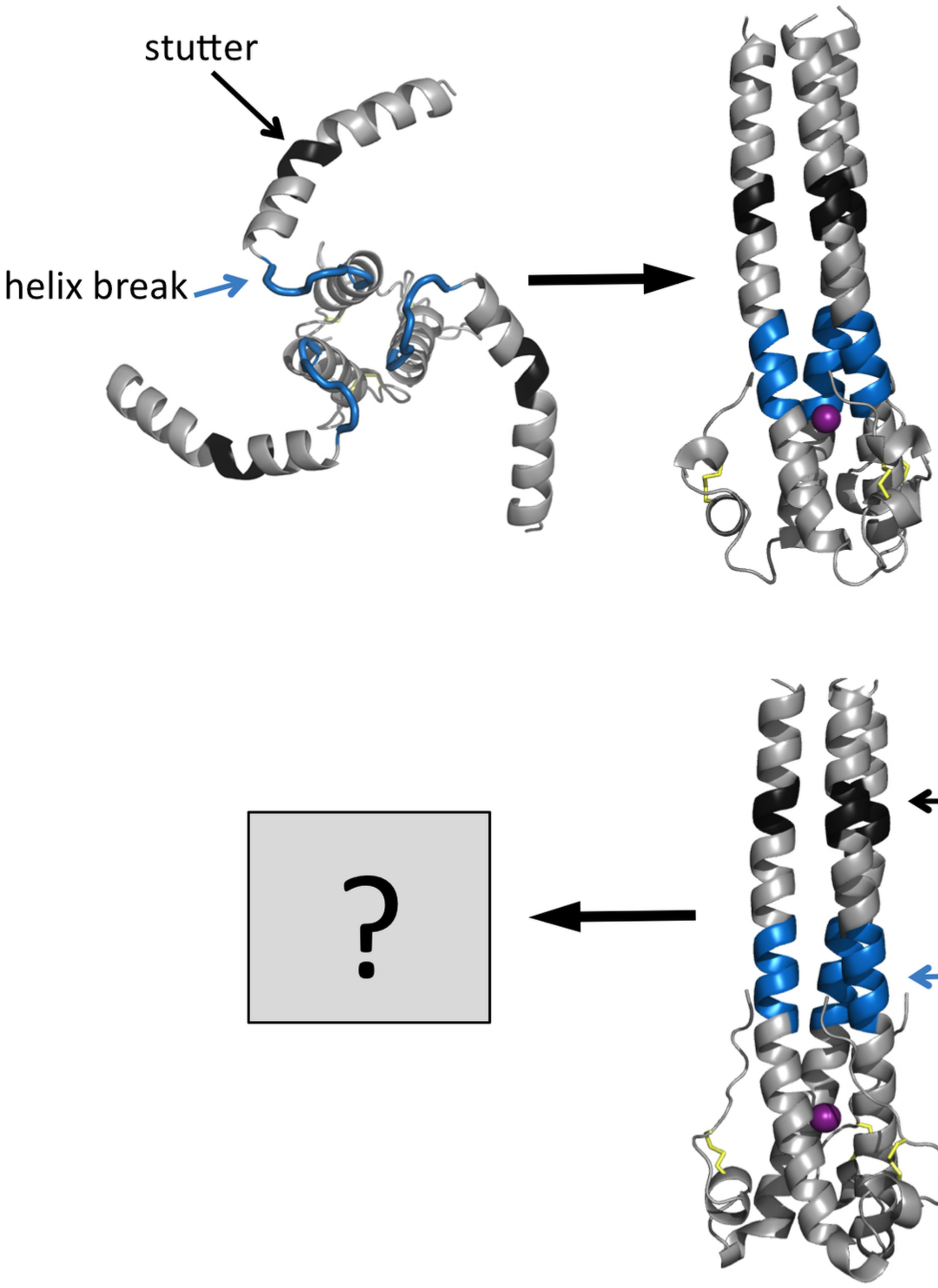

GFP = fused cells
=
Macro
STSyn1 (A555)
Nuclei (Hoechst)
Naive 293T (GFP1-10+11)
STSyn1-WT
STSyn1-C186S
STSyn1-R314S-R317S
STSyn1-L368P
STSyn1-N382A
STSyn1-N382L
STSyn1-N382D
Scale bar: 200 μm


A)
Surface expression
Fusion index
120
100
80
60
40
20
0
Surface expression
Fusion index
Naive
STSyn1-WT
STSyn1-C186S
STSyn1-R314S-R317S
STSyn1-Y347P
STSyn1-L350P
STSyn1-L354P
STSyn1-M358P
STSyn1-L365P
STSyn1-L368P
STSyn1-L372P
STSyn1-L375P
STSyn1-V379P
STSyn1-N382A
STSyn1-N382L
STSyn1-N382D
Controls
Pro variants
Asn variants
B)
Syn1-TME
EBOV GP2
MARV GP2

Pre-fusion

Post-fusion

Ebola virus GP2



Proline scanning

Syncytin-1 TM