

# Human endogenous retrovirus K (HERV-K) envelope structures in pre- and post-fusion by Cryo-EM

## Authors:

Jeremy Shek<sup>1,2,#</sup>, Chen Sun<sup>1,#</sup>, Elise M. Wilson<sup>1,2</sup>, Fatemeh Moadab<sup>3</sup>, Kathryn M. Hastie<sup>1</sup>, Roshan R. Rajamanickam<sup>1</sup>, Patrick J. Penalosa<sup>1,2</sup>, Stephanie S. Harkins<sup>1</sup>, Diptiben Parekh<sup>1</sup>, Chitra Hariharan<sup>1</sup>, Dawid S. Zyla<sup>1</sup>, Cassandra Yu<sup>1</sup>, Kelly C.L. Shaffer<sup>1,2</sup>, Victoria I. Lewis<sup>2</sup>, Ruben Diaz Avalos<sup>1</sup>, Tomas Mustelin<sup>3</sup>, and Erica Ollmann Saphire<sup>1,2\*</sup>

## Affiliations:

<sup>1</sup>Center for Vaccine Innovation, La Jolla Institute for Immunology, La Jolla, CA 92037, USA

<sup>2</sup>Dept. of Medicine, University of California San Diego. La Jolla, California 92037, USA

<sup>3</sup>Division of Rheumatology, University of Washington, Seattle, WA 98109, USA

<sup>#</sup>These authors contributed equally

\*Correspondence should be addressed to: [erica@lji.org](mailto:erica@lji.org)

## Abstract

The HERV-K envelope glycoprotein (Env) is aberrantly expressed in diseases including cancers and autoimmune disorders and is targeted by antibodies. The lack of structural information has hindered functional and immune recognition studies. We solved structures of the HERV-K Env in both pre- and post-fusion states with novel monoclonal antibodies using cryo-EM. The pre-fusion Env assembles as a trimer with a distinct fold and architecture compared to other retroviruses, while the post-fusion conformation features a unique “tether” helix within the TM subunit. A panel of monoclonal antibodies, elicited to facilitate structure determination, have been characterized for conformational and subunit specificity, serving as valuable research tools. These findings establish a structural framework for mechanistic studies of HERV-K Env in diseases and evaluation as a potential therapeutic target.

## Introduction

Over millions of years, retroviruses have infected humans and our primate ancestors, integrating into our DNA and passing down through generations. Records of these ancient viral infections are now scattered throughout the human genome, known as human endogenous retroviruses (HERVs). Genomic analyses have revealed that HERV sequences comprise 8% of the human genome, with approximately 98,000 annotated insertions(1). HERVs possess a canonical retroviral genome like their exogenous counterparts with *gag*, *pol*, and *env* genes flanked by long-terminal repeats (LTRs) containing regulatory elements(2, 3). Over time, most HERV elements accumulated mutations or deletions that rendered them non-functional and incapable of producing infectious virions(3).

Among these HERVs, the HERV-K (HML-2) family is the most recently acquired insertion; HERV-K insertions have accumulated the least number of deleterious mutations and are the most transcriptionally active. There are ~100 HERV-K full proviral inserts in our genomes, many of which contain complete open reading frames (ORFs) for each of the viral proteins(4–6). While epigenetic modifications and transcriptional repressors such as CpG methylation, histone acetylation, and chromatin remodeling keep HERV-K transcriptionally silenced in healthy adult cells(7, 8), recent studies reveal that in certain disease states, HERV-K proteins escape suppression and are actively transcribed(9–11).

Specifically, expression of the HERV-K envelope glycoprotein (Env) protein, encoded by the *env* gene, has been detected in breast, ovarian, prostate, melanoma, leukemia, and other cancers(12–21). Aberrant Env expression is associated with cancer cell proliferation, tumorigenesis, and disease progression(22–27). Autoimmune diseases including rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), and type I diabetes (T1D) have also been linked with increased expression of Env(5, 28–33). Furthermore, it is evident that the immune system develops an antibody response against Env in many cases and that antibodies against Env may even contribute to the pathogenesis of these malignancies(5, 13, 31, 34–37). Although much remains to be discovered about roles of Env in disease, the aberrant expression of Env on the cell surface could be leveraged for immunotherapy and diagnostics. Indeed, antibodies targeting HERV-K Env have demonstrated antitumor effects against breast cancer(38), engineered chimeric antigen receptor (CAR) T cells targeting Env can inhibit tumor metastasis(39, 40), and mice immunized with a vaccine directed against Env can reject tumors expressing HERV-K Env(41). Further, antibodies targeting HERV-K Env neutralize Env-mediated neurotoxicity in the cerebral spinal fluid of amyotrophic lateral sclerosis patients (ALS) by blocking binding of Env to its receptor(33, 37, 38). Collectively, these findings suggest Env as a marker or direct therapeutic target in a range of diseases.

Confounding molecular analysis of HERV-K Env function, recognition, and targeting is the lack of any structural information of the HERV-K antigen. While every human bears over ten complete copies of HERV-K *env* with the capacity for expression in their genomes, we do not yet know what the protein looks like or how antibodies may target it. In this study, we use stabilizing approaches, antibody discovery, and cryogenic electron microscopy (cryo-EM) to reveal the first structures of the HERV-K envelope glycoprotein ectodomain in both the pre-fusion and post-fusion conformations. Our findings report an overall architecture that is distinct from all known retroviral structures. We also illuminate surfaces available for immune recognition and monoclonal antibodies that begin to chart the antigenic landscape of HERV-K Env. The stabilized protein and mAbs open doors for functional studies of Env in disease and structure-based development of new interventions.

## Results

### Anti-HERV-K Env monoclonal antibody discovery

The lack of research reagents for studying HERV-K Env and its subunits hindered our initial study of the protein. The commercially available monoclonal antibodies (mAbs) HERM 1811-5

and HERM 1821-5 (Austral Biologicals) were elicited by immunization with *E. coli*-produced glycoproteins. They recognize linear epitopes and therefore cannot discern properly folded from misfolded protein and were not suitable for structural studies of the HERV-K Env.

In this study, we immunized BALB/c mice with the wild-type Env ectodomain (Env<sub>Ecto</sub>) produced in *Drosophila* S2 insect cells. Using the Bruker Beacon antibody discovery platform, followed by biochemical characterization, we selected a subset of ten monoclonal antibodies from the panel that would together broadly cover the range of conformational (pre- and post-fusion) and subunit (SU and TM) specificities (fig. 1a). We screened antibodies on ELISA coated with either purified wild-type (WT) unstabilized Env<sub>Ecto</sub>, which contains both SU and TM<sub>Ecto</sub> in a mixture of pre- and post-fusion conformations, or with post-fusion TM<sub>Ecto</sub> alone. Six of the mAbs are bound to both WT Env<sub>Ecto</sub> and post-fusion TM<sub>Ecto</sub> while the remaining four are only bound to WT Env<sub>Ecto</sub> (fig. S1a). To determine whether these mAbs recognize linear or conformational epitopes, we denatured WT Env<sub>Ecto</sub> or post-fusion TM<sub>Ecto</sub> at 95°C with DTT for 10 minutes and screened the mAbs using ELISA again. Six mAbs bound to either denatured Env<sub>Ecto</sub> or post-fusion TM<sub>Ecto</sub>, revealing that they likely bind to linearized epitopes. Six mAbs bound to either denatured Env<sub>Ecto</sub> or post-fusion TM<sub>Ecto</sub>, revealing that they likely bind to linearized epitopes. Kenv-6 only binds WT Env, and fails to bind denatured or postfusion Env<sub>Ecto</sub>, indicating it is pre-fusion specific and recognizes a conformational epitope. In contrast, while Kenv-4 binds more strongly to post-fusion TM<sub>Ecto</sub> (fig. S1a). Western blot analysis using WT Env<sub>Ecto</sub> was performed with each mAb to further confirm subunit and conformational specificities (fig. S1b). Two mAbs, Kenv-2 and Kenv-3, did not bind to fully denatured protein but did recognize Env in western blots.

Together, we found that three mAbs are SU specific (Kenv-6, Kenv-10, Kenv-14) and seven are TM specific (Kenv-1, Kenv-2, Kenv-3, Kenv-4, Kenv-7, Kenv-12). Two antibodies, Kenv-6 and Kenv-4, bind to conformational epitopes on the properly folded SU and post-fusion TM, respectively (Figure 1a). With these novel monoclonal antibodies, we could reliably detect each Env subunit when engineering Env to stabilize the pre-fusion conformation. Further, the two conformational mAbs facilitated the structure determination of Env in pre- and post-fusion states.

We sought to further validate antibody recognition to see if these mAbs would recognize HERV-K Env expressed in human patient samples as well as recombinant protein. We found that five of these antibodies, Kenv-1, Kenv-2, Kenv-3, Kenv-6, and Kenv-7 are able to stain isolated neutrophils from both SLE and RA patients (fig. 1b), which have been previously reported to express Envs from two HERV-K loci: HERV-K102 and K108 (5, 30, 31, 35). In contrast, these mAbs did not have any level of detectable antibody binding to neutrophils from healthy patients.

## HERV-K Env Pre-fusion Trimer Stabilization

There are multiple insertions of the human endogenous retrovirus K (HERV-K) in the human genome. Among them, at least 10 loci contain intact open reading frames for the *env* gene(5). These HERV-K family *env* genes are classified by the absence (type 1) or presence (type 2) of a 292-bp sequence at the 5' end of the gene, which encodes a signal peptide for plasma membrane surface localization(42). Previous studies have generated consensus sequences of

the HERV-K proviral genome by comparing HERV-K proviral inserts with the type 2 *envs*. These consensus sequences generate infectious virions(43, 44), are fusogenically active, and have broad tropism(45). In this study, we used the type 2 *Phoenix* consensus Env ectodomain (residues 97-632) sequence(43) due to >97% sequence identity to all Envs with the potential for expression (fig. S2a).

HERV-K Env is a class I viral fusion protein comprising an expected receptor-binding/surface protein subunit (SU) and a fusion/transmembrane protein subunit (TM). Three non-covalently bound SU-plus-TM protomers trimerize on the cell surface(46). Consistent with the metastability of class I viral fusion proteins, wild-type Env can spontaneously adopt the more energetically stable post-fusion conformation, in which the SU subunits disassociate, leaving the TM subunits to fold into a six-helix bundle (6HB)(47, 48). This metastability poses a significant obstacle to the expression and stability of pre-fusion conformation Env protein that would be amenable to structure determination by cryo-EM.

To overcome the problems associated with glycoprotein metastability, we explored methods that have been previously successful in the engineering of other viral envelope glycoproteins, introducing functional mutations to confer stability to the pre-fusion conformation. Modifications we designed include (i) helix-breaking proline mutations, (ii) inter-protomer disulfide linkages, and (iii) additional multimerization domains(49–51). We screened and tested mutants for expression, post-translational processing, and oligomerization, employing an iterative process of AlphaFold- and DeepCoil-driven predictions(52, 53) to determine where to place helix-breaking mutations to break the long heptad repeat 1 (HR1) helix that would otherwise drive formation of the post-fusion six-helix bundle. 22 mutants with one or two prolines within the HR1 sequence were evaluated for expression and oligomerization. The best-behaved sequences were then input into AlphaFold to generate new model predictions that were more “pre-fusion-like”. From these models, we next designed 13 potential cysteine pairs using structural visualization in ChimeraX(54, 55) and the Disulfide by Design 2.0 server(54, 55) to form disulfide bonds between SU and TM<sub>Ecto</sub>. In parallel, we assessed the utility of heterologous oligomerization domains and modified linker lengths, and evaluated expression and post-translational processing in different eukaryotic cell lines.

Among the hundreds of mutations and modifications we screened, the combination of one introduced cysteine pair between SU (V437C) and TM<sub>ecto</sub> (V498C), two point mutations that altered the furin cleavage motif from RSKR to RRRR, and addition of the T4 Fibrin trimerization domain at the C terminus of TM<sub>Ecto</sub> together resulted in the expression of stable, pre-fusion Env<sub>Ecto</sub> trimers (fig. 2a, fig. S3). Size exclusion chromatography and SDS-PAGE analysis of Env<sub>Ecto</sub> expressed in *Drosophila* S2 insect cells indicate that Env<sub>Ecto</sub> is efficiently processed by furin into SU and TM<sub>Ecto</sub> subunits which are joined by a disulfide bond, and that this Env oligomerizes as a stable trimer in solution (fig. S3b-e). We assessed these Env<sub>Ecto</sub> constructs using nano differential scanning fluorimetry and demonstrated an increase in melting temperatures from 46.0°C in the unstabilized to 53.8°C in the stabilized Env (fig. S3f-h).

Env<sub>Ecto</sub> trimers produced and purified in this way were complexed with Fab fragments from the novel Kenv-6 mAb for cryo-EM. 479,625 particles from 13,141 movies were processed with CryoSPARC(56) to obtain a reconstruction at a 2.2 Å global resolution (fig. 2b).

### **The architecture of the pre-fusion HERV-K Env<sub>Ecto</sub> Trimer**

With the high-resolution map we obtained, we were able to build nearly the entirety of Env<sub>Ecto</sub> in the pre-fusion conformation with only residues at the N- and C-terminal ends of the SU (97-99, 460-465) and the N-terminal, membrane-proximal end of TM<sub>Ecto</sub> (621-632) unmodeled (fig. 2b,c). Overall, the soluble HERV-K Env<sub>Ecto</sub> pre-fusion trimer adopts an elongated shape akin to an inverted tripod with a height of ~117 Å along the vertical axis and a width of ~83 Å. In the trimer, the SU subunits sit above their TM counterparts and are shaped like prongs that extend 49 Å from the highest point of the TM<sub>Ecto</sub> near the trimer axis (fig. 2c). At the trimer apex, between the assembled three SUs, is an open central cavity with a solvent-accessible surface area of approximately 7,000 Å<sup>2</sup>. Though the furin cleavage motif (residues 461-465) is unmodeled, the C terminus of SU ending just before the furin motif at residue 459 folds as a downward-pointing alpha-helix (α4) toward the membrane surface, and forms a bundle with the C-terminal helices (α12) of the TM<sub>Ecto</sub> (fig. 2c,d). The distance between the two residues flanking the furin sequence (V459 on the SU and F466 on TM) is 53 Å. If uncleaved, this seven amino-acid span could reach ~25 Å. A separation of 53 Å indicates that upon cleavage, the newly formed termini must displace from each other, with the newly formed SU C terminus moving toward the foot of the trimer while the newly formed TM<sub>Ecto</sub> N terminus burying itself toward the central coiled-coils. This ~30 Å displacement explains why attempts to engineer an uncleavable Env<sub>Ecto</sub> did not result in successful expression or trimerization, and instead resulted in misfolded proteins with improper quaternary structure.

### Structure of the Surface Protein Subunit (SU)

The SU (residues 97-465) is a 369 amino-acid polypeptide comprising 4 α helices and 25 β strands. The SU consists of two main subdomains, here termed the core and the base (fig. 2d).

The core domain, formed by residues 172-378, is membrane-distal at the trimer apex, and is likely involved in receptor binding to heparin, CD98HC(33, 57), and perhaps other factors yet to be discovered. The most prominent feature of the SU core is a β sandwich in which one side, comprised of five anti-parallel strands (β6, β7, β14, β17, and β18), forms an elongated, distorted sheet, while the opposing sheet is formed by two short anti-parallel strands (β13 and β15). Two disulfide bridges stabilize key loops in the core structure. The first, C275-C282, stabilizes a surface-exposed loop at the apex of the SU. This loop may be important for receptor engagement: an antibody targeting the adjacent β14 strand blocks the interaction of Env with CD98HC(33), an integrin-associated protein that mediates integrin-dependent signals that promote tumorigenesis(58). The second disulfide, C227-C246, anchors the loop flanked by the two cysteine residues, the largest and most flexible loop in our model. At higher map thresholds (e.g., 11σ), the volume for the residues 233-239 of this loop are unresolvable. Applying a blur factor of +50Å<sup>2</sup> revealed continuous density, which allowed us to trace and model the loop in its entirety.



The SU base domain, formed by residues 97-171 plus 379-439, is the more membrane-proximal of the core-base pair and is responsible for mediating interactions with the TM subunit in the pre-fusion conformation. The base consists of a distorted  $\beta$  barrel structure ( $\beta 2$ ,  $\beta 4$ ,  $\beta 5$ ,  $\beta 20$ ,  $\beta 21$ ,  $\beta 22$ , and  $\beta 24$ ) positioned above the  $\beta 1$ - $\beta 25$  sheet where the N- and C termini join. This  $\beta$  sheet extends downward toward the membrane. The C-terminal end of the SU is proximal to the very bottom of the ectodomain, where  $\alpha 4$  is located, directly preceding the furin cleavage motif (residues 461-465), which is not resolved. The SU subunit does not form any contacts with neighboring SU subunits in the trimer in our static structure and does not appear to contribute to trimeric interactions.

### Structure of the Pre-Fusion Transmembrane Protein Subunit (TM)

The transmembrane subunit (TM) is responsible for the membrane fusion mechanism of retroviral Env. The pre-fusion TM ectodomain structure presented here comprises eight  $\alpha$  helices and two  $\beta$  strands (fig. 2e). Consistent with class I viral fusion proteins, TM<sub>Ecto</sub> contains a hydrophobic fusion peptide followed by heptad repeat sequence motifs, HR1 and HR2. The fusion peptide (residues 466-486) begins as an unstructured coil that transitions into  $\alpha 5$ , sitting above  $\alpha 6$  of HR1. The first heptad repeat motif, HR1 (498-552), is divided into three  $\alpha$  helices:  $\alpha 6$ ,  $\alpha 7$ , and  $\alpha 8$ . The third,  $\alpha 8$ , is amphipathic. The hydrophobic side of the  $\alpha 8$  helix faces the center of the trimer axis and forms the central coil. HR1 and HR2 are linked by a 32-residue tether region that contains a CX<sub>7</sub>C disulfide motif (C551-C559) specific to betaretroviral envelope proteins (46). HR2 (residues 585-603) extends proximally toward the membrane, with the C-terminus of TM<sub>Ecto</sub> forming a bundle at the bottom of the trimer (Fig 2c,e).

### The Pre-Fusion TM Wraps Around SU Base

Within each protomer, the SU contacts its TM partner via the 7-stranded  $\beta$  barrel and the terminal  $\beta 1$ - $\beta 25$  sheet of the base domain. The pre-fusion TM<sub>Ecto</sub> encircles the extended N- and C-terminal strands of the base, forming a C-shaped clasp (fig. 3b,c). Helices  $\alpha 6$  and  $\alpha 7$  partially envelop the extended N- and C-terminal strands of SU before reversing direction and extending behind where  $\beta 27$  hydrogen bonds to  $\beta 25$ , joining the  $\beta 1$ - $\beta 25$  sheet in an anti-parallel fashion (fig. 3d). The TM completes its wrapping around the SU with  $\alpha 9$  enclosing the remaining portion of the clamp. Sidechain nitrogen atoms from R576 and W572 form a hydrogen bond network with the backbone oxygens of H492 and S493, effectively closing and locking the clasp (fig. 3b). As a result, the HERV-K TM forms a near-complete ring underneath the SU base around the terminal  $\beta$  sheet (fig. 3c).

Dissociation of the receptor-binding subunit from the fusion protein subunit is a trigger of conformational change to the post-fusion state(47, 48). To prevent the dissociation of the non-covalently bound SU and TM subunits, part of our previously described engineering was the introduction of an intersubunit disulfide between residues 437 on SU and 498 on TM (fig. 3e). The engineered disulfide anchors the N terminus of the TM, between the fusion peptide and HR1 at the membrane-proximal end of strand  $\beta 25$ , to the base of SU. This disulfide stabilizes the pre-fusion conformation of Env by preventing the HR1 refolding and extension that would be needed for transition to the post-fusion conformation (fig. S7b).

## TM-trimer interaction

Much like other class I viral fusion proteins such as the HIV Env, Ebola GP, RSV F, and Influenza hemagglutinin (HA) proteins, the Env trimer is stabilized by the presence of a central coiled-coil(47, 59–62). In HERV-K Env, the coiled-coil is formed by helix  $\alpha 8$  (residues 529-548) in the HR1 sequence of each TM subunit, via hydrophobic interactions between the a and d positions of the heptad repeats at the central axis (fig. 4f). We found that both the length of the  $\alpha 8$  helix as well as the *absence* of helix-breaking mutations are crucial for Env trimerization. We initially followed the HIV-1 SOSIP strategy(63) of inserting kinks within the junction between  $\alpha 7$  and  $\alpha 8$  of Env, mutating residues to prolines, and did notice an increase in expression with three such mutants. However, our best-expressing mutant, L529P, yields Env monomers alone. Even if fused to a T4 Fibrin trimerization domain, L529P Env remains monomeric and does not trimerize. We hypothesize that the proline insertion destabilizes the secondary structure of the N terminus of  $\alpha 8$ , decreasing the number of hydrophobic interactions in the central coiled-coil (fig. 4f), and thereby weakening the trimer complex. Trimerization of Env<sub>Ecto</sub> was achieved by reverting P529 back to its native lysine.

In the HERV-K Env trimer, the TM-TM interaction is further reinforced by hydrogen bonds and salt bridges interfacing HR2  $\alpha 10$  with the adjacent TM fusion peptide and HR1  $\alpha 6$  (fig. 4g). This network ensures that the outer edges of the pre-fusion TMs are fastened together in the trimeric complex.

## Glycosylation of Env

The HERV-K Env<sub>Ecto</sub> sequence contains 10 possible N-glycosylation sites on each monomer, six on the SU and four on the TM (table S2). All 10 potential glycosylation sites are occupied and we were able to build the N-acetylglucosamine (NAG) cores for all the glycans (fig. S4a). For the glycan at N128, we could also model a mannose residue on each branch (fig. S4b). For the N566 glycan, we observe strong density corresponding to an  $\alpha 1$ -6 linked fucose (fig. S4c) which we modeled. Using the GlycoSHIELD pipeline(64), we modeled all glycosylation sites except for N566 as mannose-5. For each N566 site, we used a mannose-5 with an  $\alpha 1$ -6 linked fucose. Solvent-accessible surface area analysis of the glycosylated protein using a probe radius of 1.4 Å indicates that 31% of the surface of SU is shielded compared to the unglycosylated protein, with the apex particularly solvent-exposed. The TM is more shielded with 57% of its surface covered by glycans (fig. S4d,e). Together, the cryo-EM data indicate that HERV-K Env<sub>Ecto</sub> has a 1:53 ratio of glycans to amino acids, and is less shielded than the envelope ectodomains of HIV-1 (1:23), SIV (1:27), and Jaagskiete Sheep Retrovirus (JSRV) (1:49), but more than Syncytin-1 from HERV-W (1:60) and Syncytin-2 from HERV-FRD (1:50)(65–67). Because HERV-K is typically silenced and integrated ubiquitously in all our cells, it is likely not subject to the same immune pressure as other exogenously circulating retroviruses.

## HERV-K Env adopts a novel fold

This structure of the HERV-K Envelope glycoprotein is the first pre-fusion trimer complex structure of any human endogenous retrovirus. The only other full endogenous retroviral Env structure yet available is from a hookworm endogenous viral element, a type II structure similar to flavivirus envelope proteins(68). Further, the HERV-K envelope structure presented here is

also the first pre-fusion structure available for any betaretrovirus, whether endogenous or exogenous. Among all retroviruses, pre-fusion Env trimer structures are only thus far solved for the lentiviruses HIV-1 and SIV (59, 69). For HERV-K Env, the only similarity in overall architecture is that in all three of HERV-K, HIV-1 and SIV Env, the SU domain is above (membrane distal) and the TM is below (membrane proximal), and that in all three retroviral trimers, the SU (or gp120) subunits are held in the pre-fusion trimer by the TM (or gp41), using a clasp or collar structure formed by the  $\alpha$  helices from the HR1 and HR2 domains (fig. 3b,c). Beyond those two simple organizational similarities, the overall forms, secondary structural elements, and folds differ extensively between HERV-K Env and the HIV-1/SIV lentiviruses.

Comparison of the HIV-1 and SIV envelope trimer ectodomains thus far resolved to this structure of the HERV-K Env<sub>Ecto</sub> reveals that HERV-K Env adopts a more streamlined form, primarily shaped by the SU domain. Compared to Env, the HIV and SIV trimers are shorter and wider, with heights ranging from 95 Å to 103 Å and widths between 108 Å to 120 Å (fig. 4a) vs. 117 Å tall and 83 Å wide for HERV-K Env.

The envelope glycoprotein of HIV-1 is highly dynamic, spontaneously transitioning between multiple conformational states which influence binding of antibodies, CD4, and co-receptors (70). Compared to each of the three described HIV-1 Env conformations, our structure of HERV-K Env<sub>Ecto</sub> is more like the closed, unliganded conformation of HIV-1 gp140 (state-1; PDB:5CEZ), but the global root mean squared deviation (RMSD) is very high at 23 Å (fig. 4a). State-2 (PDB:7TFO), which is the occluded-open conformation triggered by CD4 binding, differs from HERV-K Env by an RMSD of 27 Å, and state-3 (PDB:5VN3), the CD4-bound, V3 loop exposed, open conformation, differs from this HERV-K Env structure by an RMSD of 38 Å. For HIV-1, conformational dynamics are required for receptor and coreceptor binding. In contrast, in HERV-K Env, the presumed receptor binding site is exposed at the trimer apex. During cryo-EM data processing, we did observe independent movement of each SU away from the vertical central axis, causing asymmetry (fig. S11a). The appearance of asymmetry in 3D classes suggests that some “breathing” flexibility of the SU domains may occur in HERV-K Env, potentially allowing adoption of more open or closed states.

The individual SU domain of HERV-K Env has a novel fold, as revealed by a systematic structural fold search utilizing Foldseek, DALI, and HHpred servers (71–73). The global RMSDs for HERV-K SU with any of HIV-1 gp120, SIV gp120, or Syncytin-2 SU are each >23 Å, and no other protein in the databases aligns at all with the HERV-K SU (fig. 4b). The area of greatest structural homology is the base domain of HERV-K, a region that is functionally analogous to the inner domain of HIV-1/SIV gp120. Both the HERV-K base domain and the HIV/SIV inner domain contain a distorted  $\beta$  barrel and a membrane-proximal  $\beta$  sheet (fig. 4b), and both are responsible for TM/gp41 interactions. Outside the base/inner domain, the SU structures are quite different. The HERV-K SU is vertically elongated and is composed primarily of  $\beta$  strands, with a discernable, 5-stranded  $\beta$  sheet in the core. In contrast, HIV-1 and SIV gp120s are a mixture of  $\beta$  sheets and  $\alpha$  helices. Further, the lentiviral gp120s are larger than HERV-K SU (~470-510 vs. 369 amino acids), and bear both an outer domain and V1V2 variable regions, while HERV-K is a single domain structure. The V1V2 loops of HIV/SIV are variable in



sequence and critical for immune evasion. In contrast, the HERV-K sequence is highly conserved, likely a consequence of its endogenous nature and absence of immune-driven selective pressure (fig. S2). The conservation of all HERV-K Envs, however, is an advantage for the development of diagnostics and immunotherapeutics.

The human Syncytin-2 protein is critical for syncytiotrophoblast formation during placental development and is a co-opted envelope protein from the gammaretroviral HERV-FRD, inserted into the genome over 40 million years ago(74). Though the core domain of the Syncytin-2 SU subunit has been solved (PDB:7OIX), its full SU has not yet been resolved, either alone or in a trimeric complex with its TM. The Syncytin-2 SU core does, however, appear to share the  $\beta$  sheet formed by five antiparallel  $\beta$  strands extending upwards (fig. 4b). In Syncytin-2, the short loops that connect the antiparallel strands of this sheet form the binding site for MFSD2A, a membrane-embedded lysophosphatidylcholine lipid symporter(75), but it is not yet known whether the equivalent loops in HERV-K Env play a role in receptor or cofactor binding.

While the HERV-K SU is distinct from the lentiviruses, the TM<sub>Ecto</sub> subunit in its prefusion conformation exhibits a fold that is more structurally similar to that of pre-fusion HIV-1 and SIV gp41, with global RMSDs ranging from 7-10 Å (fig. 4c). The TM/gp41 of all three viral envelopes share similar tertiary structure, starting with the N-terminal fusion peptide that then becomes the C-shaped clasp predominantly composed of  $\alpha$  helices which wrap around the SU/gp120 terminal  $\beta$  sheets. A long, vertically aligned  $\alpha$  helix, which in HERV-K TM is  $\alpha$ 8 in the HR1 domain, is also present in HIV-1 and SIV gp41, and in each trimer, forms the central coiled-coil crucial for envelope trimerization (fig. 4c). This central coil structural feature is common among class I fusion proteins, even beyond retroviruses(48).

## The Post-Fusion Structure of HERV-K Env TM

Given the metastability of viral glycoproteins, both pre-fusion and post-fusion conformations of HERV-K Env are likely present on cells expressing Env. Thus, we sought to also solve the structure of the HERV-K TM ectodomain in its post-fusion conformation. The TM<sub>Ecto</sub> (residues 498-603) was expressed with a maltose-binding protein (MBP) tag connected to TM<sub>Ecto</sub> by a GS linker, a human rhinovirus (HRV) 3C protease cleavage site, and a T4 fibrin trimerization domain on the N terminus (fig. 5a). Residues 466-497 and 604-632 were not expressed due to the hydrophobicity of the fusion peptide and predicted flexibility. The purified TM<sub>Ecto</sub> fusion protein was first complexed with Fab fragments of the Kenv-4 mAb, then treated with 3C protease to remove the MBP, which prevented aggregation of TM<sub>Ecto</sub> alone (fig. 5a; fig. S5). The TM-Kenv-4 complex was purified by SEC and 696,758 particles from 19,225 movies were yield a 2.8 Å reconstruction by cryo-EM (fig. 5b; fig. S8). With the final map, we were able to model residues 502-548 and 567-603 (fig. 5c).

The HERV-K TM ectodomain in the post-fusion conformation consists of heptad repeat 1 (HR1), an internal 'tether' region (includes the <sup>551</sup>CX<sub>7</sub>C<sup>559</sup> motif), and heptad repeat 2 (HR2) (fig. 5a,d). These three features assemble into a six-helix bundle (6HB), a structural hallmark of the post-fusion conformation in type I viral glycoproteins (fig. 5c,d). At the center of the 6HB, the HR1

helices form the central trimeric coiled coil, while the HR2 and tether helices adopt an antiparallel orientation, packing against the outer surface of the HR1 (fig. 5c).

The HR1 helix features the repeating seven amino-acid sequence pattern with hydrophobic residues occupying the a and d positions, facilitating inter-helical, hydrophobic packing interactions within the trimer (fig. S6). A four-residue stutter (<sup>522</sup>LANQ<sup>525</sup>) disrupts the regular heptad pattern in the HR1 region (fig. 5d), a feature missing from the HIV-1 and SIV lentivirus gp41, but observed in Influenza hemagglutinin (HA), Ebola Virus GP, Lymphocytic Choriomeningitis Virus GP, and endogenous retroviral envelopes: Syncytin-1 and Syncytin-2 (76–79). This stutter causes deviations from the ideal helix packing geometry, leading to localized unwinding of the coiled-coil. This has been thought to impart flexibility to the helices and the trimeric core (76, 80).

## The pre- to post-fusion transition

HERV-K Env-mediated fusion requires proteolytic processing of SU and TM, a low pH environment, and dissociation of SU from TM (81). The dissociation of SU from TM initiates the irreversible springing to the low energy state, where the fusion peptide inserts into the target membrane, starting the process of membrane fusion. Alignment of the pre-fusion TM subunit to the post-fusion TM revealed two major structural rearrangements that occur around  $\alpha 8$ , which remains largely unaltered between the two conformations. The first is the formation of the elongated HR1 helix (residues 502-546) by the displacement of the fusion peptide, pulling  $\alpha 6$  and  $\alpha 7$  away from the cell membrane, which are disjointed helices forming one arm of the C-shaped clamp in the pre-fusion conformation (fig. S7b). The HR1 central helix, initially ~30 Å in length, extends to ~67 Å. This is accompanied by the second rearrangement where the TM reverses direction. The  $\alpha 9$  and  $\alpha 10$  helices pack tightly along the now extended HR1 helix, forming the tether and HR2 helices, respectively, in the post-fusion structure (fig. S7c). This reorganization positions the sequence following HR2, including the unresolved membrane-spanning region at the C terminus, closer to the target membrane, driving membrane fusion (fig. S7c,d). Residues 546-566, which include an intramolecular disulfide bond between <sup>551</sup>CX<sub>7</sub>C<sup>559</sup> and  $\beta 27$ , refold to facilitate the chain reversal. We could model this region in the pre-fusion structure, but could not resolve it in the post-fusion structure, likely due to inherent flexibility.

## Comparison to other retrovirus postfusion structures

We manually curated nine other retroviral envelope post-fusion structures from the Protein Data Bank (PDB) [HIV-1 (PDB:1I5X), SIV (PDB:1QBZ), ASLV (Avian sarcoma leukemia virus) (PDB:4JPR), HTLV-1 (Human T-lymphotropic virus 1) (PDB:1MG1), PERV (Porcine endogenous retrovirus) (PDB:7S94), Syncytin-1 (HERV-W) (PDB:6RX1), Syncytin-2 (HERV-FRD) (PDB:6RX3), XMRV (Xenotropic murine leukemia virus-related virus) (PDB:4JGS), MPMV (Mason-Pfizer monkey virus) (PDB:4JF3)] and performed sequence and structural alignments (fig. S8). All structures feature a long  $\alpha$  helix corresponding to the HR1 region (fig. S8b,c). However, while other retroviruses possess only two helices (HR1 and HR2), the post-fusion HERV-K TM core exhibits an additional helix within the tether sequence, between HR1 and HR2, which we have termed the tether helix (fig. S8b,c). Interestingly, the HERV-K tether is

closer in sequence to other retroviruses' HR2 N termini, but it does not adopt the typical HR2 position (fig. S8b). Then HERV-K TM HR2 sequence aligns with the C termini of other HR2 sequences. Though we did not model the entire HERV-K HR2, the region that was modeled (residues 585-603) do pack similarly to the compared structures on the outside of HR1. RMSDs of the HERV-K TM with other retroviral TM (ranging from ~4-20 Å for all atom pairs) are provided in table S3.

## Antibody binding interfaces

### The Kenv-6 Epitope

Three Kenv-6 Fabs were solved in complex with the trimeric prefusion Env<sub>Ecto</sub>. Kenv-6 binds to a conformational epitope at the apex of the SU, and each copy of Kenv-6 buries 569 Å<sup>2</sup> by the heavy chain and 363 Å<sup>2</sup> by the light chain. The Kenv-6 heavy chain contacts Env primarily via the complementarity determining region (CDR)H3, which forms hydrogen bonds with both side chain and backbone atoms of residues in the loops of SU proceeding β13 (residues 277-281) and between β14-β15 (residues 309-313) and β17-β18 (residues 361-364) (fig. S9a). The CDRH1 and CDRH2 of Kenv-6 form hydrogen bonds to R280 and S281. Light chain CDRL1 and CDRL2 hydrogen bond to the hydroxyl oxygens on the sidechains of both Y367 and E314 and the backbone oxygen of R363 (fig. S9b). Notably, the Kenv-6 epitope is broadly conserved among intact HERV-K Envs encoded by the human genome (fig. S2), positioning Kenv-6 as a useful research tool for the detection of Env expression and folding. Indeed, we found that Kenv-6 was able to strongly stain neutrophils isolated from an SLE patient and an RA patient (fig. 1b). Neutrophils have been reported to express Envs from both type 1 (K102) and type 2 (K108) HERV-K insertions in these diseases (30, 31).

### The Kenv-4 Epitope

Three copies of the Kenv-4 Fab are symmetrically bound to the post-fusion HERV-K TM<sub>Ecto</sub> trimer. The Kenv-4 epitope is located near the flexible loop (residues 581-585) of the TM, with each Fab spanning two adjacent monomers in their post-fusion conformation. The heavy chain of Kenv-4 buries 522 Å<sup>2</sup> of surface area on the post fusion TM, while the light chain buries 386 Å<sup>2</sup>. The Kenv-4 heavy chain interacts with two adjacent monomers, while its light chain interacts with only one monomer. The heavy chain interaction is mediated by CDRH1 and CDRH2 (fig. S10a). Here, residue D32 of the CDRH1 salt bridges to R539 on one monomer (Prot<sub>A</sub>) and H571 on the adjoining monomer (Prot<sub>B</sub>); in the complex three copies of Kenv-4 make three such interactions around the trimer. Light chain binding is only mediated by CDRL1 with six hydrogen bonds, and Y70 in framework 3 hydrogen binding to E583 (fig. S10a). The post-fusion conformation-specific Kenv-4 did not react with neutrophils from either of the patient samples evaluated here, although other TM-binding mAbs did (fig. 1b).

## Discussion

HERVs have been present for millions of years and throughout human evolution. Although each one of us carries nearly 100 copies of the HERV-K (HML-2) provirus in our own genome,

surprisingly, science thus far knows relatively little about the structures, functions and activities of these genes and their gene products when expressed. HERVs were previously thought to be silent “junk” DNA, but an emerging body of evidence now demonstrates activation of HERV elements during a variety of human disease states, including cancers and autoimmune diseases, and that patients of these diseases make antibodies against Env proteins. Among all HERV insertions, the HERV-K (HML-2) family stands out for its more recent integration into the human genome. As a result, many of the HERV-K insertions retain complete open reading frames and are able to express functional copies of the viral proteins(4, 6, 10).

Of the different HERV-K antigens, the envelope glycoprotein has been a particular focus as it is expressed on the surface of human cells and elicits human antibodies. Further, RNA transcripts and protein expression of Env have been demonstrated in various cancer tissues, including but not limited to breast cancer, ovarian cancer, leukemia, and melanoma(13, 15, 19, 20, 82, 83). Expression of HERV-K Env is also associated with rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), type I diabetes, multiple sclerosis, and various other autoimmune and neurodegenerative disorders(9, 28, 31, 84, 85).

The expressed Env has functions of its own: HERV-K Env binds to and co-immunoprecipitates with CD98HC (SLC3A2), an extracellular membrane-bound protein involved in amino acid transport, integrin signaling, and tumorigenesis. Binding of HERV-K Env to CD98HC activates the mTOR pathway, which regulates cell growth, proliferation, and tumor metabolism(86, 87). Further, HERV-K Env expressed by the K108 locus on human chromosome 7 is a fully functional and fusogenic copy(88): expressed in multinucleated melanoma cells, HERV-K108 Env has been shown to mediate melanoma cell-cell fusion *in vitro*(18).

Expression of HERV-K Env proteins induces a humoral immune response against Env in cancer(13, 15, 89). Anti-Env autoantibodies are also elicited in various autoimmune disorders and may correlate with disease progression(5, 31, 34–36, 90). Immune complexes formed between Env and these antibodies can activate immune cells and pro-inflammatory responses, contributing to the pathogenesis of diseases where anti-Env immunoglobins are prevalent (5).

Despite their appearance and apparent targetability in multiple human disease states, no structures of any HERV envelope trimer have been determined. Without such a template, we lack the means to understand how HERV-K Env interacts with cellular factors to trigger fusion or downstream cell signaling, how Env is recognized by human antibodies, if such antibodies can inhibit Env functions, how the immune response against Env may exacerbate disease pathology or confer protection, or how we may guide the design of targeted therapeutic interventions and Env-based diagnostics.

In this work, we raised a panel of novel monoclonal antibodies against HERV-K Env and used them to elucidate the first cryo-EM structures of the pre-fusion HERV-K Env (SU+TM) trimer ectodomain (2.2 Å resolution) and the post-fusion TM ectodomain (2.8 Å resolution). Though the conformations of Env on cell surfaces during disease states are not yet characterized, one may expect from our understanding of the metastability of viral glycoproteins that both pre- and post-fusion structures may exist on the cell surface and both be relevant for antibody recognition.

Kenv-6, which targets the folded SU subunit with preferential binding of the pre-fusion conformation, stains neutrophils from both RA and SLE. This indicates that SU is folding properly and is present in autoimmune disease-relevant cells. In these cells, the post-fusion-specific Kenv-4 did not stain the cells although other TM-specific antibodies did. Staining may be different in cancer cells than autoimmune cells. The tumor microenvironment is characterized by more acidic pH due to the accumulation of metabolic waste(91, 92); recognition of the the acidic-pH-driven(57, 81) post-fusion conformation of TM in Env-expressing cancers may be relevant in these contexts.

The overall structure of the pre-fusion HERV-K Env diverges from that of any retroviral envelope protein solved previously, with at least 23 Å RMSD from other known structures. The HERV-K Env trimer adopts a narrower profile than its lentiviral counterparts in HIV-1 and SIV and heavily incorporates  $\beta$  strands rather than  $\alpha$  helices into the SU subunit. The 5-stranded  $\beta$  sheet at the core of the HERV-K Env SU, however, is also present in the SU core domain of Syncytin-2, a former endogenous retrovirus protein co-opted for human placental function(75). In the TM subunit of HERV-K, the core molecular machinery that guides 6HB formation and membrane fusion is similar to gp41 from HIV-1 and SIV. Other features, however, are unique, including a CX<sub>7</sub>C sequence motif found in betaretroviral envelopes(46) and the “tether”  $\alpha$  helix in the post-fusion conformation. Further, HERV-K Env is not as glycosylated as exogenous retroviral envelopes (23 vs. 53 amino acids per glycan) and our analysis of its glycan shield reveals surfaces of vulnerability, particularly in SU, for targeting by anti-Env antibodies. These differences underscore the evolutionary divergence of HERV-K from other families of retroviruses.

The structures, stabilizing mutations, and antibodies presented here offer valuable templates and essential research tools that open doors for detection of HERV-K expression and analysis of its role in human disease states. HERV-K Env expression is observed in autoimmunity and five anti-Env antibodies described here stain human neutrophils from autoimmune disease patient samples, but not healthy neutrophils. In cancer, HERV-K Env has garnered significant interest as a possible neo-antigen or tumor-associated antigen(21, 93). Indeed, strategies targeting Env for immunotherapy have been met with promising results: an anti-Env mAb termed 6H5 has been shown to inhibit growth and induce apoptosis of breast cancer cells *in vitro* in a dose-dependent manner. Used *in vivo*, mice bearing xenograft tumors displayed significantly reduced growth with 6H5 compared to an IgG control(38). Further, a nanobody targeting Env was reported to induce antibody-dependent cellular cytotoxicity in lymphoma cell lines and can be internalized by Env-expressing cells, warranting further exploration into anti-Env antibody-drug conjugates(94). Moreover, CAR-T cells engineered to target HERV-K Env have been effective in lysing breast cancer and melanoma tumors, inhibiting cell proliferation, and preventing metastasis *in vitro* and *in vivo*(39, 40). The GNK-301 mAb (GeNeuro) binds to a linear epitope located on the  $\beta$ 14 strand at top of Env SU, blocks its interaction with CD98HC, and mitigates Env-related neurotoxicity in ALS(33). Our Kenv-6 mAb also binds to the SU, forming contacts with residues within the loop connecting  $\beta$ 14 to  $\beta$ 15 and the adjacent loop between  $\beta$ 17- $\beta$ 18, although it is not yet known whether it has the same functionality.



Previous investigations to determine if there are autoantibody responses against Env in autoimmune diseases and immune responses in cancer have largely utilized linear peptides of Env or bacterially expressed Env, thereby failing to capture antibodies that target conformational epitopes or intact, glycosylated Env(34, 35, 90), and were performed in the absence of structural information on Env. The structures we describe now provide the blueprints for functional exploration of HERV-K Env. Further, the engineered and wild-type antigens described can be used in these studies, as well as to discover mAbs from patient samples and determination of which Env epitopes are immunodominant on the surface of cells. The results presented here also open doors for use of specific mAbs to detect or target diseased cells, or neutralize Env-mediated effects in disease, either by outcompeting patient-derived antibodies or by blocking binding to surface receptors.

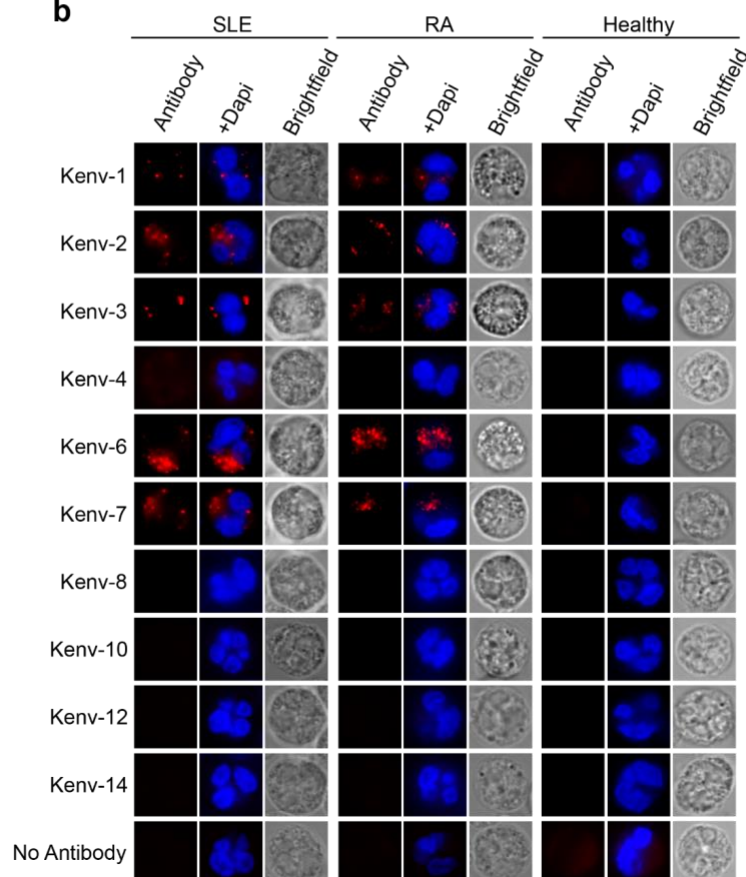
This study establishes a structural basis for exploring the function(s) of the many HERV-K *env* loci and their gene products in disease pathogenesis, and provides tools needed for development of novel immunotherapies, diagnostics, and vaccines.

# 564 Figures

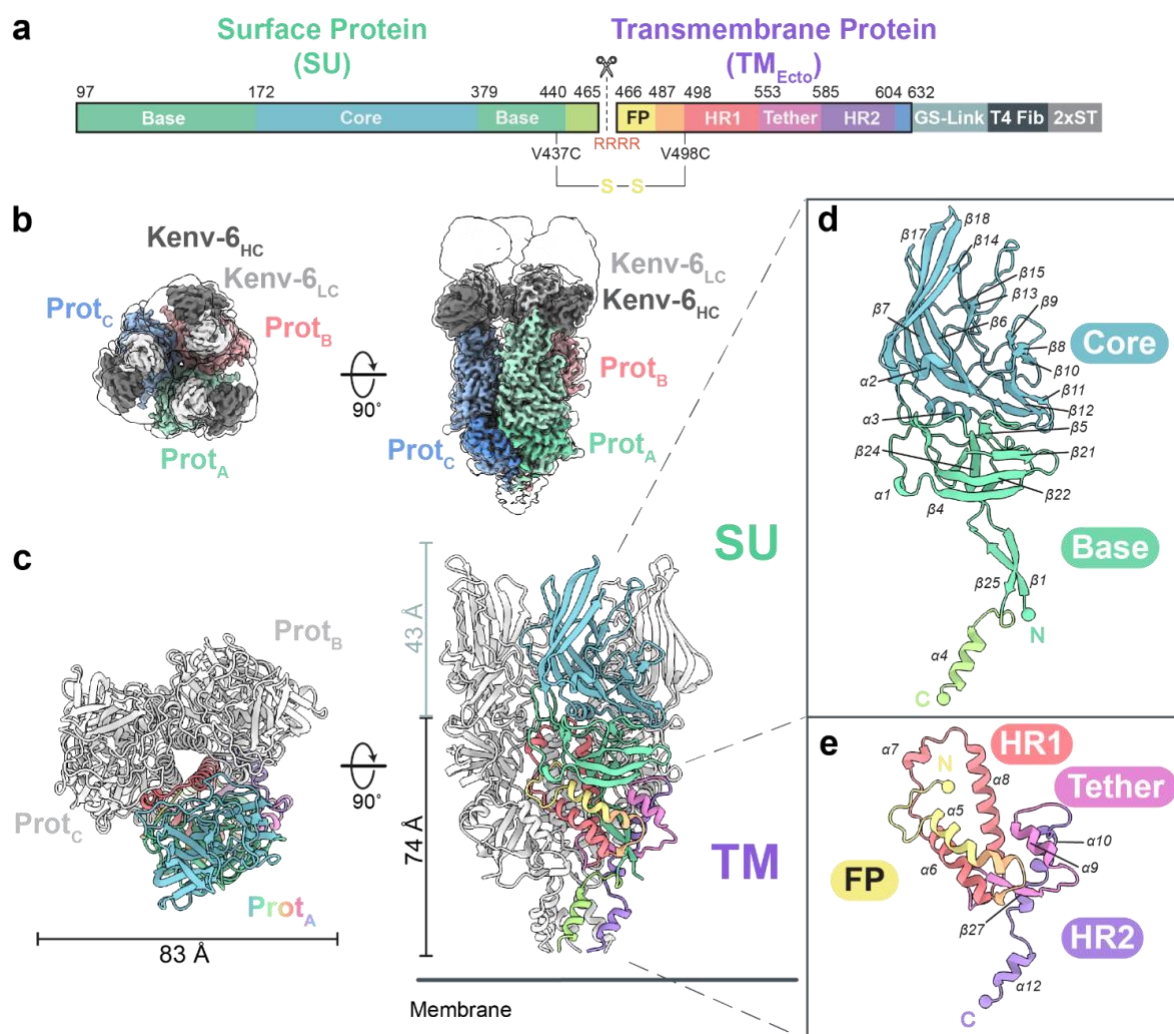
**a**

Antibody Name	Subunit	VH gene	VL gene	CDRH3 Length	Epitope Type	Western Blot
Kenv-1	TM	IGHV10-1	IGKV14-111	7	Linear	+
Kenv-2	TM	IGHV3-1	IGKV5-39	15	Linear*	+
Kenv-3	TM	IGHV8-8	IGKV10-94	17	Linear*	+
Kenv-4	TM	IGHV3-2	IGKV4-68	12	Conformational; Post-Fusion	–
Kenv-6	SU	IGHV9-3	IGKV3-5	14	Conformational; Pre-Fusion	–
Kenv-7	TM	IGHV2-6-7	IGKV12-46	11	Linear	+
Kenv-8	TM	IGHV8-11	IGKV10-94	17	Linear	+
Kenv-10	SU	IGHV1-22	IGKV9-124	5	Linear	+
Kenv-12	TM	IGHV8-8	IGKV10-94	17	Linear	+
Kenv-14	SU	IGHV3-2	IGKV8-21	9	Linear	+

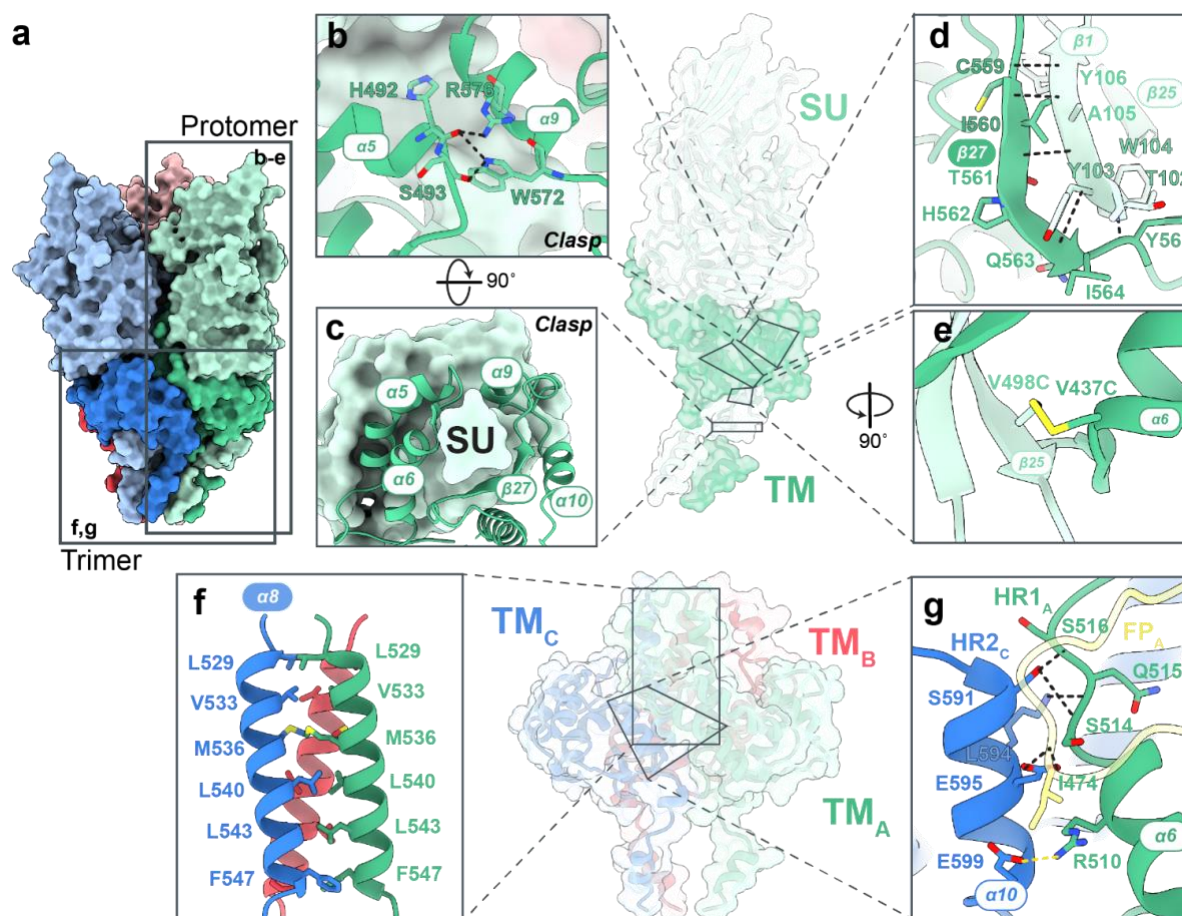
**b**



**Figure 1. Novel monoclonal antibodies against HERV-K Env. a.** Table of anti-Env mAbs discovered in this study. Antibodies were screened for surface protein (SU) or transmembrane protein subunit (TM) binding using ELISA shown in fig. S1a. Epitope type was determined using ELISAs in combination with western blots (fig. S1b). \* = presumed linear based on western blotting results. **b.** Immunofluorescence staining of isolated neutrophils from either SLE, RA, or healthy donors using mAbs from this study. Neutrophils from RA samples were treated with IFN $\alpha$  prior to fixation. Healthy neutrophils isolated with or without IFN $\alpha$  treatment show no staining by any mAb.

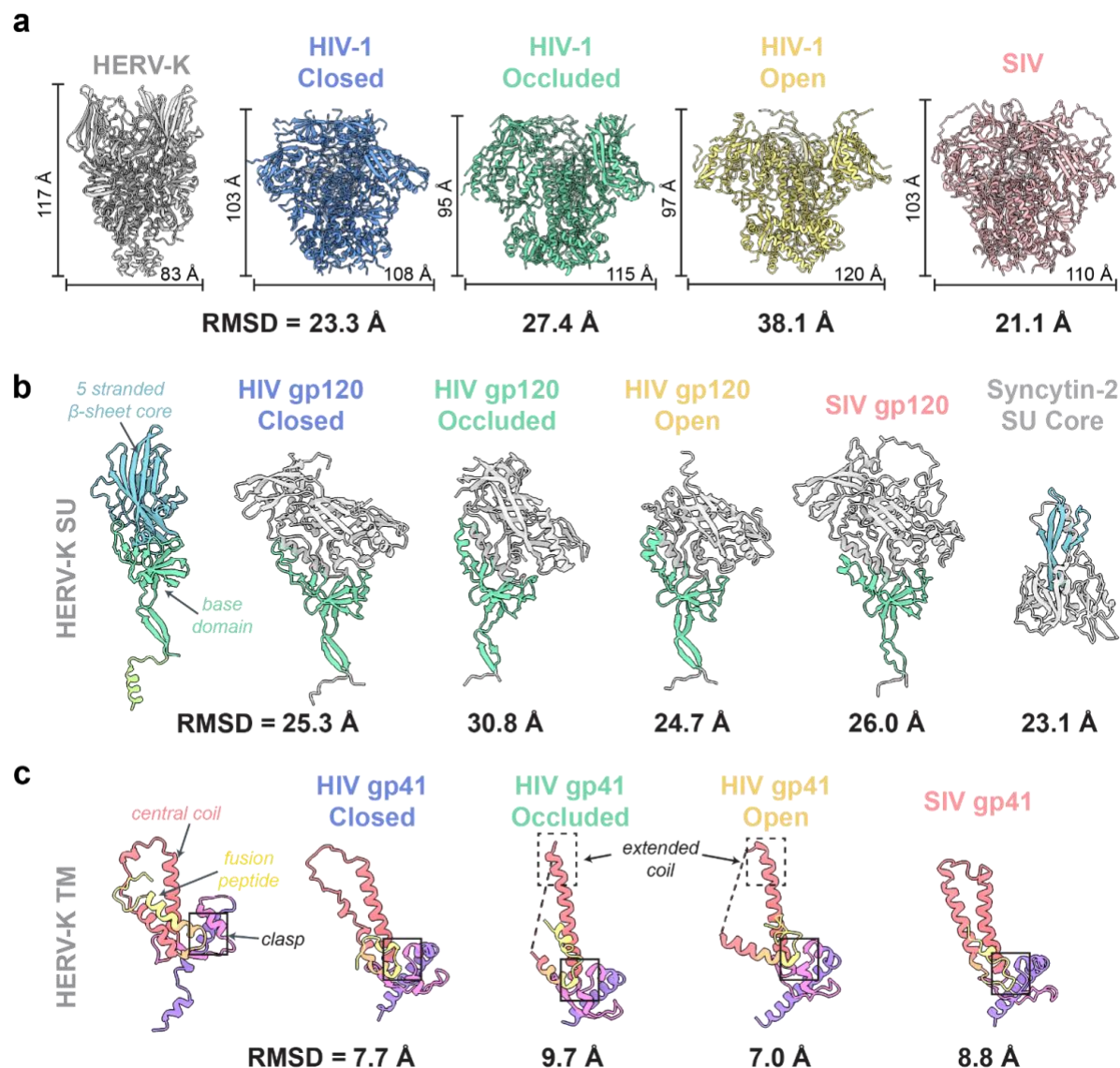


**Figure 2. Cryo-EM structure of the pre-fusion HERV-K envelope glycoprotein trimer. a.** Primary structure annotation of HERV-K Env<sub>Ecto</sub> construct. Domains are color-coded. Abbreviations: FP, fusion peptide; HR1, heptad repeat 1; HR2, heptad repeat 2; GS-Link, 12-amino acid GGS linker; T4-Fib, T4 Fibrin foldon trimerization domain; 2xST, Twin Strep-Tag II. Regions boxed in black are sequences of Env<sub>Ecto</sub>. Black scissors indicate the furin cleavage site that cleaves Env into its SU and TM subunits. Engineered disulfide bond mutations and furin-cleavage motif modifications are annotated. **b.** 2.2 Å resolution cryo-EM map of HERV-K Env<sub>Ecto</sub> trimer (colored using green, red, and blue shades) bound to three copies of the Kenv-6 Fab. Densities corresponding to the three SU subunits are colored with the lighter shade and densities for the three copies of TM are colored in the darker shade. A low-resolution map is overlaid and represented as a silhouette. **c.** Cartoon representation of the HERV-K Env<sub>Ecto</sub> trimer. Protomer A is colored in a rainbow as described in (a). The remaining 2 protomers are colored in white. **d.** Isolated cartoon representation of the surface protein (SU) subunit from an individual protomer. **e.** Cartoon representation of the transmembrane protein ectodomain (TM<sub>Ecto</sub>). Both **d** and **e** are colored using the same color scheme as (a,c). N- and C- terminal Cα atoms for each subunit are represented as spheres.



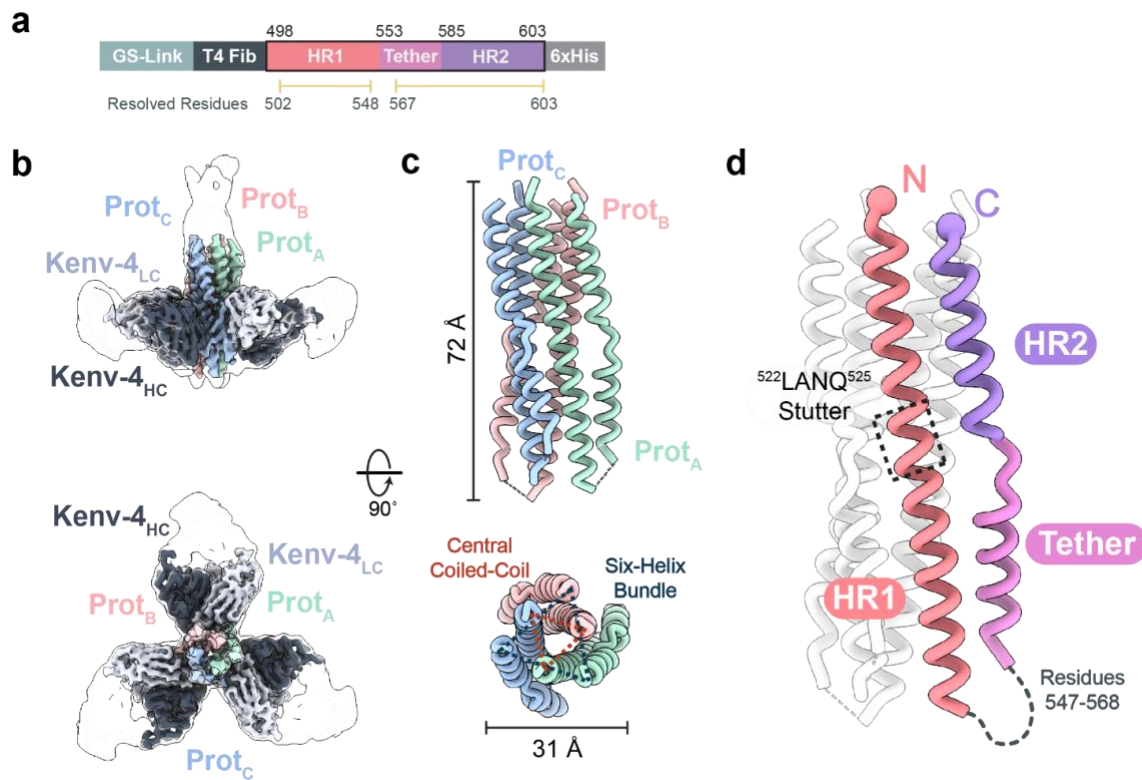
**Figure 3. Pre-fusion Env<sub>Ecto</sub> Interfaces.** **a.** The pre-fusion Env<sub>Ecto</sub> trimer displayed using a molecular surface representation. SU+TM protomer interactions are shown in panels (b-e) and trimeric interactions are shown in panels (f,g). **b.** The clasp formed by the TM around the base domain of the SU viewed from the front. Residues involved in clasp locking are shown as sticks. **c.** The C-shaped clasp formed by the TM around the SU viewed from bottom-up. **b,c.** The SU is shown as a light green surface and the TM is a dark green cartoon. **d.**  $\beta$  sheet formation between the  $\beta 1$  strand from the SU (light green) and the  $\beta 27$  strand from the TM (dark green). **e.** The engineered disulfide bond between the SU and TM to stabilize Env<sub>Ecto</sub>. **f.** Side-view of the central coiled-coil. Hydrophobic residues facing the interface from each helix are shown as sticks. **g.** The interactions between adjacent pre-fusion TM subunits distal from the central coils. The fusion peptide of TM<sub>A</sub> is colored in yellow. **b,d,g.** Hydrogen bond interactions are shown by dashed black lines and salt-bridges are shown in dashed yellow lines.





**Figure 4. Comparison of pre-fusion HERV-K Env<sub>Ecto</sub> to other solved pre-fusion retroviral structures.** **a.** Models of HERV-K Env, HIV gp (closed (PDB:5CEZ), occluded (PDB:7TFO), open (PDB:5VN3), and SIV (PDB:7T4G) ectodomain trimers. Approximate heights and widths of each complex were measured using ChimeraX (54). **b.** Comparison of receptor-binding subunits (SU: HERV-K, Syncytin-2; gp120: HIV, SIV). HERV-K Env SU is colored by subdomain while the other receptor-binding subunits are colored in light gray. Shared structural features to HERV-K SU are colored and labeled with their respective colors. **c.** Comparison of fusion protein subunits (TM: HERV-K; gp41: HIV, SIV) in pre-fusion conformation. HERV-K Env TM is colored by subdomain. Similar structural features presented on the HIV and SIV gp41 subunits are colored similarly to TM. The C-shaped clasp that encloses the SU/gp120 base domain is boxed in black on each TM/gp41. The HR1 central coil in the occluded and closed conformations of gp41 are extended at the N-terminal ends of the helix, highlighted by the dashed box. Calculated all-atom root-mean-squared-deviation (RMSD) values against the HERV-K Env<sub>Ecto</sub> complex or respective subunit are displayed under each structure.





## Acknowledgements

The authors would like to thank S. Schendel for reviewing and editing the manuscript. We acknowledge the cryo-EM facility of La Jolla Institute for Immunology for assistance in grid preparation and data collection, and equipment supported by the GHR Foundation and other generous private donors.

**Funding:** This research was supported by a Curebound Discovery Grant (13502-01-000-408) to E.O.S., research support by LJI & Kyowa Kirin, Inc. (KKNA-Kyowa Kirin North America to E.O.S. and a Kowya Kirin North America Accelerator Grant (18030-01-000-408) to J.S.).

**Author Contributions:** J.S., K.M.H., and E.O.S. conceptualized and designed experiments. J.S., E.M.W., K.M.H., S.H., and P.P. contributed to the discovery, expression, and biochemical characterization of the antibodies. F.M. and T.M. collected and imaged neutrophils. J.S., E.M.W., C.Y., and V.I.L. designed, cloned, expressed, and purified the HERV-K Env<sub>Ecto</sub> constructs. J.S., C.S., D.P., K.S., and R.R., cloned, expressed, and prepared post-fusion TM<sub>Ecto</sub> proteins. J.S., C.S., and R.D.A. collected data for single particle analysis. J.S. performed data processing, model building, and refinement of the pre-fusion Env<sub>Ecto</sub>. C.S. performed data processing, model building, and refinement of the post-fusion TM<sub>Ecto</sub>. D.Z. helped with model refinement. J.S. and C.S. prepared the figures and wrote the original draft. K.M.H. and E.O.S. reviewed and edited the manuscript. All authors provided critical feedback and helped to shape the research, analysis, and writing of the paper.

**Competing Interests:** The La Jolla Institute for Immunology has pending provisional patent coverage on the mAbs and proteins described herein. The applications, entitled "HERV-K Envelope Protein Binders and Compositions and Methods of Use Thereof" and "Stabilized Pre-Fusion HERV-K Envelope Ectodomain Trimer", lists inventors E.O.S, K.M.H, S.H., E.M.W., and J.S., and E.O.S. and J.S., respectively, and were assigned U.S. Provisional Patent Nos. 63/638,067 and 63/783,708.

**Data and materials availability:** The Kenv-6 bound pre-fusion HERV-K Env<sub>Ecto</sub> and Kenv-4 bound post-fusion TM<sub>Ecto</sub> cryo-EM maps have been deposited to the EM Database with codes EMD-0000 and EMD-0000 respectively. Atomic models are deposited in the Protein Data Bank with accession codes 0ZZZ and 0ZZZ. After publication, all research materials referenced in this study will be made available upon reasonable request through material transfer agreements.

# References

1. J. Paces, A. Pavlíček, V. Paces, HERVd: database of human endogenous retroviruses. *Nucleic Acids Res.* **30**, 205–206 (2002).
2. B. Xue, L. A. Sechi, D. J. Kelvin, Human Endogenous Retrovirus K (HML-2) in Health and Disease. *Front. Microbiol.* **11** (2020).
3. P. N. Nelson, P. R. Carnegie, J. Martin, H. Davari Ejtehad, P. Hooley, D. Roden, S. Rowland-Jones, P. Warren, J. Astley, P. G. Murray, Demystified. Human endogenous retroviruses. *Mol. Pathol.* **56**, 11–18 (2003).
4. R. P. Subramanian, J. H. Wildschutte, C. Russo, J. M. Coffin, Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology* **8**, 90 (2011).
5. M. Tokuyama, B. M. Gunn, A. Venkataraman, Y. Kong, I. Kang, T. Rakib, M. J. Townsend, K. H. Costenbader, G. Alter, A. Iwasaki, Antibodies against human endogenous retrovirus K102 envelope activate neutrophils in systemic lupus erythematosus. *J. Exp. Med.* **218** (2021).
6. M. Tokuyama, Y. Kong, E. Song, T. Jayewickreme, I. Kang, A. Iwasaki, ERVmap analysis reveals genome-wide transcription of human endogenous retroviruses. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 12565–12572 (2018).
7. L. L. K. M, M. E, M. E, M. J, CpG methylation directly regulates transcriptional activity of the human endogenous retrovirus family HERV-K(HML-2). *J. Virol.* **79**, 876–883 (2005).
8. Z. Salavatiha, R. Soleimani-Jelodar, S. Jalilvand, The role of endogenous retroviruses-K in human cancer. *Rev. Med. Virol.* **30**, e2142 (2020).
9. N. Grandi, E. Tramontano, HERV Envelope Proteins: Physiological Role and Pathogenic Potential in Cancer and Autoimmunity. *Front. Microbiol.* **0**, 462 (2018).
10. M. Garcia-Montojo, T. Doucet-O'Hare, L. Henderson, A. Nath, Human endogenous retrovirus-K (HML-2): a comprehensive review. *Crit. Rev. Microbiol.* **44**, 715–738 (2018).
11. R. F. Downey, F. J. Sullivan, F. Wang-Johanning, S. Ambs, F. J. Giles, S. A. Glynn, Human endogenous retrovirus K and cancer: Innocent bystander or tumorigenic accomplice? *Int. J. Cancer* **137**, 1249–1257 (2015).
12. R. Contreras-Galindo, M. H. Kaplan, P. Leissner, T. Verjat, I. Ferlenghi, F. Bagnoli, F. Giusti, M. H. Dosik, D. F. Hayes, S. D. Gitlin, D. M. Markovitz, Human endogenous retrovirus K (HML-2) elements in the plasma of people with lymphoma and breast cancer. *J. Virol.* **82**, 9329–9336 (2008).
13. F. Wang-Johanning, M. Li, F. J. Esteva, K. R. Hess, B. Yin, K. Rycaj, J. B. Plummer, J. G. Garza, S. Ambs, G. L. Johanning, Human endogenous retrovirus type K antibodies and mRNA as serum biomarkers of early-stage breast cancer: Endogenous retroviruses as serum breast cancer biomarkers. *Int. J. Cancer* **134**, 587–595 (2014).
14. J.-O. Jo, Y.-J. Kang, M. S. Ock, K. S. Song, M.-J. Jeong, S.-J. Jeong, Y. H. Choi, E.-J. Ko,

- S.-H. Leem, S. Kim, H.-S. Kim, H.-J. Cha, Expression profiles of HERV-K Env protein in normal and cancerous tissues. *Genes Genomics* **38**, 91–107 (2015).
15. F. Wang-Johanning, J. Liu, K. Rycaj, M. Huang, K. Tsai, D. G. Rosen, D. T. Chen, D. W. Lu, K. F. Barnhart, G. L. Johanning, Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer. *Int. J. Cancer* **120**, 81–90 (2007).
16. M. Natoli, J. Gallon, H. Lu, A. Amgheib, D. J. Pinato, F. A. Mauri, T. Marafioti, A. U. Akarca, I. Ullmo, J. Ip, E. O. Aboagye, R. Brown, A. Karadimitris, S. Ghaem-Maghami, Transcriptional analysis of multiple ovarian cancer cohorts reveals prognostic and immunomodulatory consequences of ERV expression. *J. Immunother. Cancer* **9** (2021).
17. T. A. Wallace, R. F. Downey, C. J. Seufert, A. Schetter, T. H. Dorsey, C. A. Johnson, R. Goldman, C. A. Loffredo, P. Yan, F. J. Sullivan, F. J. Giles, F. Wang-Johanning, S. Ambis, S. A. Glynn, Elevated HERV-K mRNA expression in PBMC is associated with a prostate cancer diagnosis particularly in older men and smokers. *Carcinogenesis* **35**, 2074–2083 (2014).
18. G. Huang, Z. Li, X. Wan, Y. Wang, J. Dong, Human endogenous retroviral K element encodes fusogenic activity in melanoma cells. *J. Carcinog.* **12**, 5 (2013).
19. S. Depil, C. Roche, P. Dussart, L. Prin, Expression of a human endogenous retrovirus, HERV-K, in the blood cells of leukemia patients. *Leukemia* **16**, 254–259 (2002).
20. T. Li, K. Qian, J. Han, Y. Liu, L. Jia, X. Wang, T. Li, B. Zhang, J. Li, H. Li, L. Dou, L. Li, Higher expression of human endogenous retrovirus-K was observed in peripheral B lymphocytes of leukemia and lymphoma patients. *AIDS Res. Hum. Retroviruses* **40**, 268–279 (2024).
21. C. Li, Q. Qian, C. Yan, M. Lu, L. Li, P. Li, Z. Fan, W. Lei, K. Shang, P. Wang, J. Wang, T. Lu, Y. Huang, H. Yang, H. Wei, J. Han, J. Xiao, F. Chen, H. Atlas, HervD Atlas: a curated knowledgebase of associations between human endogenous retroviruses and diseases. *Nucleic Acids Res.* **1**, 13–14 (2013).
22. F. Zhou, M. Li, Y. Wei, K. Lin, Y. Lu, J. Shen, G. L. Johanning, F. Wang-Johanning, Activation of HERV-K Env protein is essential for tumorigenesis and metastasis of breast cancer cells. *Oncotarget* **7**, 84093–84117 (2016).
23. C. Lemaître, J. Tsang, C. Bireau, T. Heidmann, M. Dewannieux, A human endogenous retrovirus-derived gene that can contribute to oncogenesis by activating the ERK pathway and inducing migration and invasion. *PLoS Pathog.* **13**, e1006451 (2017).
24. M. Li, L. Radvanyi, B. Yin, K. Rycaj, J. Li, R. Chivukula, K. Lin, Y. Lu, J. Shen, D. Z. Chang, D. Li, G. L. Johanning, F. Wang-Johanning, Downregulation of Human Endogenous Retrovirus Type K (HERV-K) Viral env RNA in Pancreatic Cancer Cells Decreases Cell Proliferation and Tumor Growth. *Clin. Cancer Res.* **23**, 5892–5911 (2017).
25. E.-J. Ko, M.-S. Ock, Y.-H. Choi, J. L. Iovanna, S. Mun, K. Han, H.-S. Kim, H.-J. Cha, Human Endogenous Retrovirus (HERV)-K env Gene Knockout Affects Tumorigenic Characteristics of nupr1 Gene in DLD-1 Colorectal Cancer Cells. *Int. J. Mol. Sci.* **22** (2021).
26. Y. Masuda, R. Ishihara, Y. Murakami, S. Watanabe, Y. Asao, N. Gotoh, T. Kasamatsu, H.

- Takei, N. Kobayashi, T. Saitoh, H. Murakami, H. Handa, Clinical significance of human endogenous retrovirus K (HERV-K) in multiple myeloma progression. *Int. J. Hematol.* **117**, 563–577 (2023).
27. J. Zhao, K. Rycaj, S. Geng, M. Li, J. B. Plummer, B. Yin, H. Liu, X. Xu, Y. Zhang, Y. Yan, S. A. Glynn, T. H. Dorsey, S. Ambs, G. L. Johanning, L. Gu, F. Wang-Johanning, Expression of Human Endogenous Retrovirus Type K Envelope Protein is a Novel Candidate Prognostic Marker for Human Breast Cancer. *Genes Cancer* **2**, 914 (2011).
28. F. Reynier, T. Verjat, F. Turrel, P. E. Imbert, H. Marotte, B. Mougin, P. Miossec, Increase in human endogenous retrovirus HERV-K (HML-2) viral load in active rheumatoid arthritis. *Scand. J. Immunol.* **70**, 295–299 (2009).
29. J. Sicut, N. Sutkowski, B. T. Huber, Expression of human endogenous retrovirus HERV-K18 superantigen is elevated in juvenile rheumatoid arthritis. *The Journal of Rheumatology* **32** (2005).
30. A. Laine, X. Wang, K. Ni, S. E. B. Smith, R. Najjar, L. S. Whitmore, M. Yacoub, A. Bays, M. Gale, T. Mustelin, Expression of Envelope Protein Encoded by Endogenous Retrovirus K102 in Rheumatoid Arthritis Neutrophils. *Microorganisms* **11**, 1310 (2023).
31. A. I. Khadjinova, X. Wang, A. Laine, K. Ukadike, M. Eckert, A. Stevens, A. A. Bengtsson, C. Lood, T. Mustelin, Autoantibodies against the envelope proteins of endogenous retroviruses K102 and K108 in patients with systemic lupus erythematosus correlate with active disease. *Clin. Exp. Rheumatol.* **40**, 1306–1312 (2022).
32. P. W. Halcrow, D. N. K. Quansah, N. Kumar, J. P. Steiner, A. Nath, J. D. Geiger, HERV-K (HML-2) envelope protein induces mitochondrial depolarization and neurotoxicity via endolysosome iron dyshomeostasis. *J. Neurosci.* **44**, e0826232024 (2024).
33. J. P. Steiner, M. Bachani, N. Malik, C. DeMarino, W. Li, K. Sampson, M. H. Lee, J. Kowalak, M. Bhaskar, T. Doucet-O'Hare, M. Garcia-Montojo, M. Cowen, B. Smith, L. B. Reoma, J. Medina, J. Brunel, J. Pierquin, B. Charvet, H. Perron, A. Nath, Human Endogenous Retrovirus K Envelope in Spinal Fluid of Amyotrophic Lateral Sclerosis Is Toxic. *Ann. Neurol.* **92**, 545–561 (2022).
34. G. Mameli, G. L. Erre, E. Caggiu, S. Mura, D. Cossu, M. Bo, M. L. Cadoni, A. Piras, N. Mundula, E. Colombo, G. Buscetta, G. Passiu, L. A. Sechi, Identification of a HERV-K env surface peptide highly recognized in Rheumatoid Arthritis (RA) patients: a cross-sectional case–control study. *Clin. Exp. Immunol.* **189**, 127 (2017).
35. X. Wang, A. Hefton, K. Ni, K. C. Ukadike, M. A. Bowen, M. Eckert, A. Stevens, C. Lood, T. Mustelin, Autoantibodies against unmodified and citrullinated human endogenous retrovirus K envelope protein in patients with rheumatoid arthritis. *J. Rheumatol.* **49**, 26–35 (2022).
36. M. Noli, G. Meloni, S. Ruberto, S. Jasemi, E. R. Simula, D. Cossu, M. Bo, M. Palermo, L. A. Sechi, HERV-K Envelope Protein Induces Long-Lasting Production of Autoantibodies in T1DM Patients at Onset in Comparison to ZNT8 Autoantibodies. *Pathogens* **11**, 1188 (2022).
37. K. W. Ng, J. Boumelha, K. S. S. Enfield, J. Almagro, H. Cha, O. Pich, T. Karasaki, D. A. Moore, R. Salgado, M. Sivakumar, G. Young, M. Molina-Arcas, S. de Carné Trécesson, P.



Anastasiou, A. Fendler, L. Au, S. T. C. Shepherd, C. Martínez-Ruiz, C. Puttick, J. R. M. Black, T. B. K. Watkins, H. Kim, S. Shim, N. Faulkner, J. Attig, S. Veeriah, N. Magno, S. Ward, A. M. Frankell, M. Al Bakir, E. L. Lim, M. S. Hill, G. A. Wilson, D. E. Cook, N. J. Birkbak, A. Behrens, N. Yousaf, S. Popat, A. Hackshaw, A. Rowan, A. Huebner, B. B. Campbell, C. Bailey, C. Lee, D. Biswas, E. Colliver, F. Athanasopoulou, H. Zhai, J. K. Rane, K. Grigoriadis, M. Dietzen, M. Leung, M. Angelova, O. Lucas, O. Al-Sawaf, R. Rosenthal, J. Nicod, A. Bunkum, A. Toncheva, C. Abbosh, C. Richard, C. Naceur-Lombardelli, F. Gimeno-Valiente, J. M. Lam, K. Thol, K. Thakkar, M. Werner Sunderland, M. D. Forster, N. Kanu, P. Prymas, R. Benthams, S. Saghafeina, S. A. Quezada, S. Vanloo, S. Zaccaria, S. M. Lee, S. Hessey, W. K. Liu, D. Papadatos-Pastos, J. Wilson, S. Benafif, T. Ahmad, E. Borg, M. Falzon, R. Khirya, T. Marafioti, A. Sharp, C. Pilotti, H. K. Dhanda, K. Chan, N. Gower, R. Leslie, S. Smith, A. G. Nicholson, E. Lim, J. Herrero, C. Castignani, E. Larose Cadieux, J. Demeulemeester, P. Van Loo, K. S. Peggs, C. Veiga, G. Royle, C.-A. Collins-Fekete, A. J. Procter, A. Nair, A. Ahmed, M. N. Taylor, N. Navani, R. M. Thakrar, D. Lawrence, D. Patrini, E. Nye, R. K. Stone, D. Chuter, M. MacKenzie, F. Fraioli, P. Ashford, S. M. Janes, M. Tanić, S. Beck, A. Rice, A. Devaraj, C. Proli, D. Kaniu, H. Bhayani, H. Chavan, H. Raubenheimer, L. Ambrose, M. Malima, N. Fernandes, P. De Sousa, P. Shah, S. Booth, S. I. Buder, S. Jordan, S. Begum, E. Boleti, A. Stewart, A. Magness, C. E. Weeden, D. Levi, E. Grönroos, J. Goldman, M. Escudero, P. Hobson, R. Vendramin, S. Boeing, T. Denner, V. Barbè, W.-T. Lu, W. Hill, Y. Naito, Z. Ramsden, A. Grapa, H. Zhang, K. AbdulJabbar, X. Pan, K. Gilbert, A. Karamani, B. Chain, D. R. Pearce, D. Karagianni, E. Hoxha, F. Gálvez-Cancino, G. Stavrou, G. Mastrokalos, H. L. Lowe, I. Matos, J. L. Reading, J. A. Hartley, K. Selvaraju, K. Chen, L. Ensell, M. Shah, M. Vasquez, M. Litovchenko, O. Chervova, P. Pawlik, R. E. Hynds, S. López, S. Gamble, S. K. A. Ung, S. K. Bola, T. P. Mourikis, V. Spanswick, Y. Wu, E. Martinoni Hoogenboom, F. Monk, J. W. Holding, J. Choudhary, K. Bhakhri, M. Scarci, M. Hayward, N. Panagiotopoulos, P. Gorman, R. C. M. Stephens, S. Bandula, Y. N. S. Wong, T. Clark, H. Cheyne, M. Khalil, S. Richardson, T. Cruickshank, B. Naidu, G. Matharu, J. A. Shaw, J. Riley, L. Primrose, J. Le Quesne, K. G. Blyth, A. Kerr, A. Clipson, A. Chaturvedi, C. Dive, D. G. Rothwell, E. Kilgour, J. Tugwood, L. Priest, P. Oliveira, P. Crosbie, G. Price, J. Cave, K. M. Kerr, C. R. Lindsay, F. H. Blackhall, M. G. Krebs, Y. Summers, A. Kirk, M. Thomas, M. Asif, N. Kostoulas, R. Bilancia, G. Middleton, M. J. Shackcloth, A. Leek, J. D. Hodgkinson, N. Totten, C. Dick, L. Robinson, P. Russell, M. Hewish, S. Danson, J. F. Lester, F. Gomes, K. Brown, M. Carter, A. J. Patel, A. Osman, C. Lacson, G. Langman, H. Shackelford, M. Djearaman, S. Kadiri, A. Alzetani, J. Richards, L. Scarlett, P. Ingram, S. Chee, S. Austin, A. Bajaj, A. Nakas, A. Sodha-Ramdeen, D. A. Fennell, K. Ang, M. Tufail, M. F. Chowdhry, M. Scotland, R. Boyles, S. Rathinam, C. Wilson, D. Marrone, S. Dulloo, A. Montero, E. Smith, E. Fontaine, F. Granato, H. Doran, J. Novasio, K. Rammohan, L. Joseph, P. Bishop, R. Shah, S. Moss, V. Joshi, H. J. W. L. Aerts, T. L. Kaufmann, R. F. Schwarz, J. Kisistok, M. Sokac, M. Diossy, Z. Szallasi, K. Dijkstra, Y. Yuan, F. Byrne, L. A. Boos, B. Shum, C. L. Gerard, A. M. Schmitt, C. Messiou, D. Cunningham, I. Chau, N. Starling, N. Turner, L. Welsh, R. L. Jones, J. Droney, S. Banerjee, K. C. Tatham, S. Jhanji, K. Harrington, A. Okines, A. Reid, K. Young, A. J. S. Furness, L. Pickering, E. Nicholson, S. Kumar, K. A. Wilkinson, A. Swerdlow, R. J. Wilkinson, C. T. Hiley, K. Litchfield, N. McGranahan, M. Jamal-Hanjani, J. Larkin, S.-H. Lee, S. Turajlic, C. Swanton, J. Downward, G. Kassiotis, Antibodies against endogenous retroviruses promote lung cancer immunotherapy. *Nature* **124**, 1–11 (2023).

38. F. Wang-Johanning, K. Rycak, J. B. Plummer, M. Li, B. Yin, K. Frerich, J. G. Garza, J. Shen, K. Lin, P. Yan, S. A. Glynn, T. H. Dorsey, K. K. Hunt, S. Ambs, G. L. Johanning, Immunotherapeutic Potential of Anti-Human Endogenous Retrovirus-K Envelope Protein Antibodies in Targeting Breast Tumors. *JNCI: Journal of the National Cancer Institute* **104**,

189–210 (2012).

39. J. Krishnamurthy, B. A. Rabinovich, T. Mi, K. C. Switzer, S. Olivares, S. N. Maiti, J. B. Plummer, H. Singh, P. R. Kumaresan, H. M. Huls, F. Wang-Johanning, L. J. N. Cooper, Genetic engineering of T cells to target HERV-K, an ancient retrovirus on melanoma. *Clin. Cancer Res.* **21**, 3241–3251 (2015).
40. F. Zhou, J. Krishnamurthy, Y. Wei, M. Li, K. Hunt, G. L. Johanning, L. J. N. Cooper, F. Wang-Johanning, Chimeric antigen receptor T cells targeting HERV-K inhibit breast cancer and its metastasis through downregulation of Ras. <https://doi.org/10.1080/2162402X.2015.1047582> **4** (2015).
41. B. Kraus, K. Fischer, S. M. Büchner, W. S. Wels, R. Löwer, K. Sliva, B. S. Schnierle, Vaccination Directed against the Human Endogenous Retrovirus-K Envelope Protein Inhibits Tumor Growth in a Murine Model System. *PLoS One* **8**, e72756 (2013).
42. A. Ruggieri, E. Maldener, M. Sauter, N. Mueller-Lantzsch, E. Meese, O. T. Fackler, J. Mayer, Human endogenous retrovirus HERV-K(HML-2) encodes a stable signal peptide with biological properties distinct from Rec. *Retrovirology* **6**, 17 (2009).
43. M. Dewannieux, F. Harper, A. Richaud, C. Letzelter, D. Ribet, G. Pierron, T. Heidmann, Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. *Genome Res.* **16**, 1548 (2006).
44. Y. N. Lee, P. D. Bieniasz, Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog.* **3**, e10 (2007).
45. P. Kramer, V. Lausch, A. Volkwein, K. Hanke, O. Hohn, N. Bannert, The human endogenous retrovirus K(HML-2) has a broad envelope-mediated cellular tropism and is prone to inhibition at a post-entry, pre-integration step. *Virology* **487**, 121–128 (2016).
46. J. E. Henzy, J. M. Coffin, Betaretroviral Envelope Subunits Are Noncovalently Associated and Restricted to the Mammalian Class. *J. Virol.* **87**, 1937 (2013).
47. P. M. Colman, M. C. Lawrence, The structural biology of type I viral membrane fusion. *Nat. Rev. Mol. Cell Biol.* **4**, 309–319 (2003).
48. J. M. White, S. E. Delos, M. Brecher, K. Schornberg, Structures and Mechanisms of Viral Membrane Fusion Proteins: Multiple Variations on a Common Theme. *Crit. Rev. Biochem. Mol. Biol.* **43**, 189–219 (2008).
49. H. Ebel, T. Benecke, B. Vollmer, Stabilisation of Viral Membrane Fusion Proteins in Prefusion Conformation by Structure-Based Design for Structure Determination and Vaccine Development. MDPI [Preprint] (2022). <https://doi.org/10.3390/v14081816>.
50. X. Yang, L. Florin, M. Farzan, P. Kolchinsky, P. D. Kwong, J. Sodroski, R. Wyatt, Modifications That Stabilize Human Immunodeficiency Virus Envelope Glycoprotein Trimers in Solution. *J. Virol.* **74**, 4746 (2000).
51. B. Pedenko, G. Sulbaran, D. Guilligay, G. Effantin, W. Weissenhorn, SARS-CoV-2 S Glycoprotein Stabilization Strategies. *Viruses* **15** (2023).
52. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K.

- 871 Tunyasuvunakool, R. Bates, A. Židek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl,  
872 A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S.  
873 Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T.  
874 Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli,  
875 D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–  
876 589 (2021).
- 877 53. J. Ludwiczak, A. Winski, K. Szczepaniak, V. Alva, S. Dunin-Horkawicz, DeepCoil-a fast and  
878 accurate prediction of coiled-coil domains in protein sequences. *Bioinformatics* **35**, 2790–  
879 2795 (2019).
- 880 54. E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H.  
881 Morris, T. E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators,  
882 and developers. *Protein Sci.* **30**, 70 (2021).
- 883 55. D. B. Craig, A. A. Dombkowski, Disulfide by Design 2.0: A web-based tool for disulfide  
884 engineering in proteins. *BMC Bioinformatics* **14**, 1–7 (2013).
- 885 56. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid  
886 unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290–296 (2017).
- 887 57. L. R. Robinson-McCarthy, K. R. McCarthy, M. Raaben, S. Piccinotti, J. Nieuwenhuis, S. H.  
888 Stubbs, M. J. G. Bakkers, S. P. J. Whelan, Reconstruction of the cell entry pathway of an  
889 extinct virus. *PLoS Pathog.* **14**, e1007123 (2018).
- 890 58. C. C. Feral, N. Nishiya, C. A. Fenczik, H. Stuhlmann, M. Slepak, M. H. Ginsberg, CD98hc  
891 (SLC3A2) mediates integrin signaling. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 355–360 (2005).
- 892 59. D. Lyumkis, J.-P. Julien, N. de Val, A. Cupo, C. S. Potter, P.-J. Klasse, D. R. Burton, R. W.  
893 Sanders, J. P. Moore, B. Carragher, I. A. Wilson, A. B. Ward, Cryo-EM structure of a fully  
894 glycosylated soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1484–1490 (2013).
- 895 60. J. E. Lee, M. L. Fusco, A. J. Hessel, W. B. Oswald, D. R. Burton, E. O. Saphire, Structure  
896 of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* **454**,  
897 177–182 (2008).
- 898 61. J. S. McLellan, M. Chen, S. Leung, K. W. Graepel, X. Du, Y. Yang, T. Zhou, U. Baxa, E.  
899 Yasuda, T. Beaumont, A. Kumar, K. Modjarrad, Z. Zheng, M. Zhao, N. Xia, P. D. Kwong, B.  
900 S. Graham, Structure of RSV Fusion Glycoprotein Trimer Bound to a Prefusion-Specific  
901 Neutralizing Antibody. *Science* **340**, 1113 (2013).
- 902 62. I. A. Wilson, J. J. Skehel, D. C. Wiley, Structure of the haemagglutinin membrane  
903 glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366–373 (1981).
- 904 63. D. Wrapp, Z. Mu, B. Thakur, K. Janowska, O. Ajayi, M. Barr, R. Parks, K. Mansouri, R. J.  
905 Edwards, B. H. Hahn, P. Acharya, K. O. Saunders, B. F. Haynes, Structure-Based  
906 Stabilization of SOSIP Env Enhances Recombinant Ectodomain Durability and Yield. *J Virol*  
907 **97**, e0167322 (2023).
- 908 64. Y.-X. Tsai, N.-E. Chang, K. Reuter, H.-T. Chang, T.-J. Yang, S. von Bülow, V. Sehrawat, N.  
909 Zerrouki, M. Tuffery, M. Gecht, I. L. Grothaus, L. Colombi Ciacchi, Y.-S. Wang, M.-F. Hsu,  
910 K.-H. Khoo, G. Hummer, S.-T. D. Hsu, C. Hanus, M. Sikora, Rapid simulation of

glycoprotein structures by grafting and steric exclusion of glycan conformer libraries. *Cell* **187**, 1296–1311.e26 (2024).

65. M. Gholami Barzoki, S. Shatizadeh Malekshahi, Z. Heydarifard, M. J. Mahmodi, H. Soltanghoreae, The important biological roles of Syncytin-1 of human endogenous retrovirus W (HERV-W) and Syncytin-2 of HERV-FRD in the human placenta development. *Mol Biol Rep* **50**, 7901–7907 (2023).

66. P. Priščáková, M. Svoboda, Z. Feketová, J. Hutník, V. Repiská, H. Gbelcová, L. Gergely, Syncytin-1, syncytin-2 and suppressyn in human health and disease. *J Mol Med (Berl)* **101**, 1527–1542 (2023).

67. A. Dupressoir, C. Lavialle, T. Heidmann, From ancestral infectious retroviruses to bona fide cellular genes: role of the captured syncytins in placentation. *Placenta* **33**, 663–671 (2012).

68. M. Merchant, C. P. Mata, Y. Liu, H. Zhai, A. V. Protasio, Y. Modis, A bioactive phlebovirus-like envelope protein in a hookworm endogenous virus. *Sci. Adv.* **8**, eabj6894 (2022).

69. R. Andrabi, J. Pallesen, J. D. Allen, G. Song, J. Zhang, N. de Val, G. Gegg, K. Porter, C.-Y. Su, M. Pauthner, A. Newman, H. Bouton-Verville, F. Garces, I. A. Wilson, M. Crispin, B. H. Hahn, B. F. Haynes, L. Verkoczy, A. B. Ward, D. R. Burton, The chimpanzee SIV envelope trimer: Structure and deployment as an HIV vaccine template. *Cell Rep.* **27**, 2426–2441.e6 (2019).

70. H. T. Nguyen, Q. Wang, S. Anang, J. G. Sodroski, Characterization of the human immunodeficiency virus (HIV-1) envelope glycoprotein conformational states on infectious virus particles. *J. Virol.* **97**, e0185722 (2023).

71. M. van Kempen, S. S. Kim, C. Tumescheit, M. Mirdita, J. Lee, C. L. M. Gilchrist, J. Söding, M. Steinegger, Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **42**, 243–246 (2024).

72. L. Holm, Dali server: structural unification of protein families. *Nucleic Acids Res.* **50**, W210–W215 (2022).

73. J. Söding, A. Biegert, A. N. Lupas, The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–8 (2005).

74. A. Vargas, J. Moreau, S. Landry, F. LeBellego, C. Toufaily, E. Rassart, J. Lafond, B. Barbeau, Syncytin-2 plays an important role in the fusion of human trophoblast cells. *J. Mol. Biol.* **392**, 301–318 (2009).

75. M. Martinez-Molledo, E. Nji, N. Reyes, Structural insights into the lysophospholipid brain uptake mechanism and its inhibition by syncytin-2. *Nat. Struct. Mol. Biol.* **29**, 604–612 (2022).

76. C. D. Higgins, V. N. Malashkevich, S. C. Almo, J. R. Lai, Influence of a heptad repeat stutter on the pH-dependent conformational behavior of the central coiled-coil from influenza hemagglutinin HA2: The Heptad Repeat Stutter in Influenza HA2. *Proteins* **82**, 2220–2228 (2014).

77. W. Weissenhorn, A. Carfí, K. H. Lee, J. J. Skehel, D. C. Wiley, Crystal structure of the



950 Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain.  
951 *Mol. Cell* **2**, 605–616 (1998).

952 78. S. Igonet, M.-C. Vaney, C. Vonrhein, G. Bricogne, E. A. Stura, H. Hengartner, B. Eschli, F.  
953 A. Rey, X-ray structure of the arenavirus glycoprotein GP2 in its postfusion hairpin  
954 conformation. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19967–19972 (2011).

955 79. K. Ruigrok, M.-C. Vaney, J. Buchrieser, E. Baquero, J. Hellert, B. Baron, P. England, O.  
956 Schwartz, F. A. Rey, M. Backovic, X-ray structures of the post-fusion 6-helix bundle of the  
957 human syncytins and their functional implications. *J. Mol. Biol.* **431**, 4922–4940 (2019).

958 80. J. H. Brown, C. Cohen, D. A. Parry, Heptad breaks in alpha-helical coiled coils: stutters and  
959 stammers. *Proteins* **26**, 134–145 (1996).

960 81. L. R. Robinson, S. P. J. Whelan, Infectious Entry Pathway Mediated by the Human  
961 Endogenous Retrovirus K Envelope Protein. *J Virol* **90**, 3640–3649 (2016).

962 82. A. Serafino, E. Balestrieri, P. Pierimarchi, C. Matteucci, G. Moroni, E. Oricchio, G. Rasi, A.  
963 Mastino, C. Spadafora, E. Garaci, P. S. Vallebona, The activation of human endogenous  
964 retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation. *Exp. Cell*  
965 *Res.* **315**, 849–862 (2009).

966 83. E. Stricker, E. C. Peckham-Gregory, M. E. Scheurer, HERVs and Cancer—A  
967 Comprehensive Review of the Relationship of Human Endogenous Retroviruses and  
968 Human Cancers. *Biomedicines* **11** (2023).

969 84. S. Levet, J. Medina, J. Joanou, A. Demolder, N. Queruel, K. Réant, M. Normand, M.  
970 Seffals, J. Dimier, R. Germi, T. Piofczyk, J. Portoukalian, J. L. Touraine, H. Perron, An  
971 ancestral retroviral protein identified as a therapeutic target in type-1 diabetes. *JCI Insight* **2**  
972 (2017).

973 85. V. Gröger, H. Cynis, Human endogenous retroviruses and their putative role in the  
974 development of autoimmune disorders such as multiple sclerosis. *Front. Microbiol.* **9**  
975 (2018).

976 86. T. Wang, M. Medynets, K. R. Johnson, T. T. Doucet-O'Hare, B. DiSanza, W. Li, Y. Xu, A.  
977 Bagnell, R. Tyagi, K. Sampson, N. Malik, J. Steiner, A. Hadegean, J. Kowalak, J. O'Malley,  
978 D. Maric, A. Nath, Regulation of stem cell function and neuronal differentiation by HERV-K  
979 via mTOR pathway. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 17842–17853 (2020).

980 87. Z. Zou, T. Tao, H. Li, X. Zhu, mTOR signaling pathway and mTOR inhibitors in cancer:  
981 progress and challenges. *Cell Biosci.* **10**, 31 (2020).

982 88. M. Dewannieux, S. Blaise, T. Heidmann, Identification of a Functional Envelope Protein  
983 from the HERV-K Family of Human Endogenous Retroviruses. *J. Virol.* **79**, 15573 (2005).

984 89. A. Kleiman, N. Senyuta, A. Tryakin, M. Sauter, A. Karseladze, S. Tjulandin, V. Gurtsevitch,  
985 N. Mueller-Lantzsch, HERV-K(HML-2) GAG/ENV antibodies as indicator for therapy effect  
986 in patients with germ cell tumors. *Int. J. Cancer* **110**, 459–461 (2004).

987 90. C. A. Hervé, E. B. Lugli, A. Brand, D. J. Griffiths, P. J. W. Venables, Autoantibodies to  
988 human endogenous retrovirus-K are frequently detected in health and disease and react



- 989 with multiple epitopes. *Clin. Exp. Immunol.* **128**, 75 (2002).
- 990 91. V. Estrella, T. Chen, M. Lloyd, J. Wojtkowiak, H. H. Cornnell, A. Ibrahim-Hashim, K. Bailey,  
991 Y. Balagurunathan, J. M. Rothberg, B. F. Sloane, J. Johnson, R. A. Gatenby, R. J. Gillies,  
992 Acidity generated by the tumor microenvironment drives local invasion. *Cancer Res.* **73**,  
993 1524–1535 (2013).
- 994 92. E. Boedtker, S. F. Pedersen, The acidic tumor microenvironment as a driver of cancer.  
995 *Annu. Rev. Physiol.* **82**, 103–126 (2020).
- 996 93. S. Hosseini-porham, L. A. Sechi, Anti-HERV-K drugs and vaccines, possible therapies  
997 against tumors. *Vaccines (Basel)* **11** (2023).
- 998 94. F. Spriano, L. Cascione, J. Sgrignani, N. Bendik, S. Napoli, G. Sartori, E. Cannas, T. Gong,  
999 A. J. Arribas, M. Pizzi, D. Rossi, D. F. Robbiani, A. Cavalli, F. Bertoni, Characterization of a  
1000 novel humanized heavy chain antibody targeting endogenous retroviruses with anti-  
1001 lymphoma activity, *bioRxiv* (2024). <https://doi.org/10.1101/2024.01.17.576027>.
- 1002 95. M. Silva, Y. Kato, M. B. Melo, I. Phung, B. L. Freeman, Z. Li, K. Roh, J. W. Van Wijnbergen,  
1003 H. Watkins, C. A. Enemuo, B. L. Hartwell, J. Y. H. Chang, S. Xiao, K. A. Rodrigues, K. M.  
1004 Cirelli, N. Li, S. Haupt, A. Aung, B. Cossette, W. Abraham, S. Kataria, R. Bastidas, J.  
1005 Bhiman, C. Linde, N. I. Bloom, B. Groschel, E. Georgeson, N. Phelps, A. Thomas, J. Bals,  
1006 D. G. Carnathan, D. Lingwood, D. R. Burton, G. Alter, T. P. Padera, A. M. Belcher, W. R.  
1007 Schief, G. Silvestri, R. M. Ruprecht, S. Crotty, D. J. Irvine, A particulate saponin/TLR  
1008 agonist vaccine adjuvant alters lymph flow and modulates adaptive immunity. *Sci. Immunol.*  
1009 **6**, eabf1152 (2021).
- 1010 96. Graphene Oxide Grid Preparation (2016). <https://doi.org/10.6084/m9.figshare.3178669.v1>.
- 1011 97. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L.  
1012 Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill,  
1013 D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A.  
1014 Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B.  
1015 Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S.  
1016 Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A.  
1017 Židek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J. M. Jumper, Accurate structure  
1018 prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
- 1019 98. K. Jamali, L. Käll, R. Zhang, A. Brown, D. Kimanius, S. H. W. Scheres, Automated model  
1020 building and protein identification in cryo-EM maps. *Nature* **628**, 450–457 (2024).
- 1021 99. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta*  
1022 *Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).
- 1023 100. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta*  
1024 *Crystallogr D Biol Crystallogr* **60**, 2126–2132 (2004).
- 1025 101. R. Sanchez-Garcia, J. Gomez-Blanco, A. Cuervo, J. M. Carazo, C. O. S. Sorzano, J.  
1026 Vargas, DeepEMhancer: a deep learning solution for cryo-EM volume post-processing.  
1027 *Commun Biol* **4**, 874 (2021).
- 1028 102. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze,

- 1029 L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant,  
1030 R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H.  
1031 Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams,  
1032 Macromolecular structure determination using X-rays, neutrons and electrons: recent  
1033 developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861–877 (2019).
- 1034 103. C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V.  
1035 Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall 3rd, J.  
1036 Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, D. C. Richardson, MolProbity: More  
1037 and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293–  
1038 315 (2018).
- 1039 104. UniProt Consortium, UniProt: The universal protein knowledgebase in 2023. *Nucleic*  
1040 *Acids Res.* **51**, D523–D531 (2023).
- 1041 105. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam,  
1042 M. Remmert, J. Söding, J. D. Thompson, D. G. Higgins, Fast, scalable generation of high-  
1043 quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539  
1044 (2011).
- 1045 106. P. V. Troshin, J. B. Procter, G. J. Barton, Java bioinformatics analysis web services for  
1046 multiple sequence alignment--JABAWS:MSA. *Bioinformatics* **27**, 2001–2002 (2011).

## Acknowledgements

The authors would like to thank S. Schendel for reviewing and editing the manuscript. We acknowledge the cryo-EM facility of La Jolla Institute for Immunology for assistance in grid preparation and data collection, and equipment supported by the GHR Foundation and other generous private donors.

**Funding:** This research was supported by a Curebound Discovery Grant (13502-01-000-408) to E.O.S., research support by LJI & Kyowa Kirin, Inc. (KKNA-Kyowa Kirin North America to E.O.S. and a Kowya Kirin North America Accelerator Grant (18030-01-000-408) to J.S.).

**Author Contributions:** J.S., K.M.H., and E.O.S. conceptualized and designed experiments. J.S., E.M.W., K.M.H., S.H., and P.P. contributed to the discovery, expression, and biochemical characterization of the antibodies. F.M. and T.M. collected and imaged neutrophils. J.S., E.M.W., C.Y., and V.I.L. designed, cloned, expressed, and purified the HERV-K Env<sub>Ecto</sub> constructs. J.S., C.S., D.P., K.S., and R.R., cloned, expressed, and prepared post-fusion TM<sub>Ecto</sub> proteins. J.S., C.S., and R.D.A. collected data for single particle analysis. J.S. performed data processing, model building, and refinement of the pre-fusion Env<sub>Ecto</sub>. C.S. performed data processing, model building, and refinement of the post-fusion TM<sub>Ecto</sub>. D.Z. helped with model refinement. J.S. and C.S. prepared the figures and wrote the original draft. K.M.H. and E.O.S. reviewed and edited the manuscript. All authors provided critical feedback and helped to shape the research, analysis, and writing of the paper.

**Competing Interests:** The La Jolla Institute for Immunology has pending provisional patent coverage on the mAbs described herein. The application, entitled "HERV-K Envelope Protein

1072 Binders and Compositions and Methods of Use Thereof", lists inventors E.O.S, K.M.H, S.H.,  
1073 E.M.W., and J.S., was assigned U.S. Provisional Patent No. 63/638,067.

1074

1075 **Data and materials availability:** The Kenv-6 bound pre-fusion HERV-K Env<sub>Ecto</sub> and Kenv-4  
1076 bound post-fusion TM<sub>Ecto</sub> cryo-EM maps have been deposited to the EM Database with codes  
1077 EMD-0000 and EMD-0000 respectively. Atomic models are deposited in the Protein Data Bank  
1078 with accession codes 0ZZZ and 0ZZZ. After publication, all research materials referenced in this  
1079 study will be made available upon reasonable request through material transfer agreements.