Receptor-recognition and antiviral mechanisms of retrovirus-derived human proteins

Shashank Khare, Miryam I Villalba, Juan C Canul-Tec, Arantza Balsebre Cajiao, Anand Kumar, Marija Backovic, Felix A Rey, Els Pardon, Jan Steyaert, Camilo Perez, et al.

Receptor-recognition and antiviral mechanisms of retrovirus-derived human proteins

Shashank Khare[1#], Miryam I. Villalba[1#], Juan C. Canul-Tec[1#], Arantza Balsebre Cajiao[2#], Anand Kumar[#1], Marija Backovic[3], Felix A. Rey[3], Els Pardon[4,5], Jan Steyaert[4,5], Camilo Perez[2*], Nicolas Reyes[1*]

[1]Fundamental Microbiology and Pathogenicity Unit, CNRS, Université de Bordeaux, IECB, Bordeaux, France

[2]Biozentrum, University of Basel, Basel, Switzerland

[3]Institut Pasteur, Université Paris Cité, CNRS UMR3569, Unité de Virologie Structurale, Paris, France

[4]Structural Biology Brussels, Vrije Universiteit Brussel, VUB, Brussels, Belgium

[5]VIB-VUB Center for Structural Biology, VIB, Brussels, Belgium

# Contributed equally to this work

*Correspondence: camilo.perez@unibas.ch; nicolas.reyes@u-bordeaux.fr

Human syncytin-1 and supressyn are cellular proteins of retroviral origin that catalyze and regulate cell-cell fusion events to establish the maternal-fetal interface in the placenta[1-5]. In cell culture, they restrict infections from members of the largest interference group of vertebrate retroviruses[2,6-9], including simian viruses with zoonotic potential, and are regarded as host-immunity factors expressed during development[10]. At the core of the syncytin-1 and supressyn functions are poorly understood mechanisms by which they recognize a common cellular receptor, the membrane transporter ASCT2. Here, we present cryo-electron microscopy structures of human ASCT2 in complexes with the receptor binding domains of syncytin-1 and supressyn. Despite significant evolutionary divergence, the two placental proteins engage similar binding interfaces on a large aqueous vestibule in ASCT2 within the membrane plane, stabilized by the formation of a hybrid β-sheet or "clamp" with the receptor. Structural predictions of the receptor binding domains of extant retroviruses indicate that they recognize ASCT2 through the same ancient molecular mechanism, revealing that the placental proteins compete with the retroviruses for overlapping binding interfaces and clamping sites. Together, this work uncovers a common ASCT2-recognition mechanism by a large group of endogenous and disease-causing retroviruses, and provides a high-resolution view on how retrovirus-derived human proteins exert morphological and immunological functions in placenta.

Main

The interface between the mother and fetus in the human placenta is formed by fusion of embryonic cells into a large syncytial structure known as the syncytiotrophoblast[11]. The underlying membrane fusion is catalyzed by retrovirus-derived proteins, known as syncytins[12]. Syncytin-1 is the envelope (Env) protein of the human endogenous retrovirus W (HERV-W)[13] that was endogenized millions of years ago, and has retained its membrane-fusogenic function[1,2]. It is expressed in hominoid placentae[3], and its dysregulation is associated with abnormal placentation[14-16]. Syncytin-1 is a single-pass transmembrane protein that folds into trimers expressed at the plasma membrane[17,18], and catalyzes membrane fusion upon binding to the neutral amino acid transporter ASCT2[2,19]. The mature syncytin-1 has the organization of the gamma-type Envs bearing an N-terminal surface subunit with the receptor binding domain ($SYCY1_{RBD}$) that is covalently linked to the C-terminal transmembrane subunit[20,21]. The latter anchors the trimer to the membrane and conveys the membrane fusion machinery[22]. In turn, supressyn is encoded by the *env* gene of a divergent HERV (HERV-H48), which is expressed in placenta as a soluble protein that encompasses the receptor binding domain of the ancient Env protein ($SUPYN_{RBD}$) and lacks fusogenic function. Supressyn also uses ASCT2 as the cellular receptor and acts as a negative regulator of syncytia formation mediated by syncytin-1[4,5,23].

Exogenous retroviruses of the largest superinfection interference group in vertebrates, namely the RD411-and-D-type-retrovirus (RDR) group, recognize ASCT2 orthologs at the cell surface for cell-entry to initiate the infectious cycle[6-8]. The RDR-group includes type-C feline (RD411-RV), baboon (BaEV), and avian (REV) retroviruses, as well as the type D simian retroviruses (SRV1-7)[24]. The SRVs are prevalent in non-human primates[25,26] and hold zoonotic potential based on the recently established "spillover" ranking[27], while the REVs are widely distributed among different bird species and have a harmful impact on the global poultry industry with no

69 available vaccines[28]. The Env glycoproteins of the RDR-group retroviruses recognize ASCT2
70 and are able to infect human cells. The invariant amino acid motif $SDGGGX_2DX_2R$ present in
71 both the RDR Env proteins and syncytin-1 is believed to play a key role in receptor
72 recognition[29]. Notably, supressyn lacks the above-mentioned motif and is able to inhibit
73 infection by RDR-group retroviruses in cell-culture[10] through unknown receptor interference
74 mechanisms. Thus, a general role of HERV-derived proteins, and in particular that of supressyn
75 as host-immunity restriction factors against retroviruses has been postulated[10].
76 ASCT2 is a sodium-dependent neutral amino acid transporter expressed in peripheral organs[30]
77 and an emergent drug target in cancer therapy, as it constitutes an upregulated route of
78 glutamine uptake for malignant cell proliferation[31,32]. ASCT2 is a solute carrier 1 (SLC1)
79 protein, and the reported cryo-EM structures[33,34] confirmed the structural fold characteristic of
80 the family[35,36]: ASCT2 assembles as a bowl-shaped homotrimer with a large extracellular
81 aqueous vestibule that reaches down approximately halfway across the membrane. Each
82 subunit contains two transmembrane-helical domains, the scaffold (scaD) and transport (tranD)
83 domains. The scaDs from the three subunits pack to form a central propeller-like structure that
84 anchors the protein to the membrane, while the three tranDs localize between the blades of the
85 propeller. The three tranDs work independently and, through rigid-body movements, alternate
86 between outward- and inward-facing states to translocate the substrate across the membrane in
87 an elevator-like fashion[37,38]. Early comparative studies, using ASCT2 mammalian orthologs,
88 suggested that a stretch of poorly-conserved residues on the extracellular surface of the
89 transporter acts as a negative determinant region that precludes syncytin-1 and/or retroviral
90 recognition in rodent orthologs[39]. This region maps to a β-strand hairpin (β-hairpin) on the
91 ASCT2 extracellular vestibule[33], but the molecular mechanism(s) underlying the recognition of
92 ASCT2 by endogenous or exogenous retroviral proteins has remained unknown.

In this work, we applied structural and functional approaches to determine how syncytin-1 and suppresyn recognize ASCT2, and uncovered an ancient structural mechanism that is conserved among a large group of related endogenous or exogenous retroviral proteins. We also demonstrate that $SYCY1_{RBD}$, $SUPYN_{RBD}$, and a newly discovered nanobody act as inhibitors of ASCT2 transport that could advance the development of novel selective compounds for cancer therapy.

Cryo-EM structures determination

To improve the stability and yield of recombinant proteins for *in vitro* structural and functional analyses, we exchanged residues in the sequence of the human wild-type ASCT2 ($ASCT2_{WT}$) for consensus residues of structural orthologs[40]. The consensus design ($ASCT2_{CO}$) shares ~90% amino acid identity with the $ASCT2_{WT}$, and retains the syncytin-1 receptor function for cell-cell fusion in a split-GFP complementation assay[41] (Extended Data Fig. 1a,b). Moreover, $ASCT2_{CO}$ reconstituted in proteoliposomes showed ~5-fold higher steady-state transport activity than $ASCT2_{WT}$, demonstrating that the consensus design is more resistant to inactivation in detergent (Extended Data Fig. 1c).

Supressyn and a syncytin-1 construct encompassing their respective RBDs were expressed as secreted proteins, and a biochemically well-behaved construct of syncytin-1 was obtained after a truncation at residue G439 ($SYCY1_{\Delta 439}$). The $SYCY1_{\Delta 439}$ and supressyn formed stable complexes with ASCT2 with apparent dissociation constant ($^{app}K_D$) values of 1.0 and 0.2 μM, respectively (Extended Data Fig. 2a-c), and inhibited cell-cell fusion mediated by syncytin-1 and $ASCT2_{WT}$ (Extended Data Fig. 1b). These results demonstrate that the receptor-recognition mechanisms of the recombinant $SYCY1_{\Delta 439}$ and supressyn proteins remained intact.

During cryo-EM sample optimization, we screened $ASCT2_{WT}$ and $ASCT2_{CO}$ complexes with $SYCY1_{\Delta 439}$ or supressyn in detergent solutions and obtained final cryo-EM maps of the $ASCT2_{WT}$-$SYCY1_{\Delta 439}$ and $ASCT2_{CO}$-supressyn complexes at overall resolutions of 2.6Å and

3.4Å, respectively (Extended Data Table 1, Extended Data Fig. 3-5). The cryo-EM maps enabled model building of the $SYNC1_{RBD}$ (residues 21-142) and $SUPYN_{RBD}$ (residues 61-138) bound to the receptor.

Receptor-recognition mechanism of syncytin-1

The structure of the $ASCT2_{WT}$-$SYCY1_{RBD}$ complex reveals a single molecule of the $SYCY1_{RBD}$ bound to the aqueous extracellular vestibule of the ASCT2 trimer (Fig. 1a,b), positioned deep within the membrane plane and shielded from the lipids (Fig. 1c). The $SYCY1_{RBD}$ structure differs from those of other gamma-type RBDs (including the murine[42] and feline[43] leukemia viruses, as well as the HERV-derived syncytin-2[41] and EnvP(b)1[44]), and a search of the protein data bank (PDB)[45] showed the $SYCY1_{RBD}$ represents a so far uncharacterized fold with minimal resemblance to experimentally determined structures. The novel fold includes a central core region formed by a three-stranded anti-parallel β-sheet (β1-β3) (Fig. 1d). The core β-sheet is flanked by a partially unstructured region and a short alpha-helix (α_N) on the N-terminal side, and by a long alpha-helix (α_C) at the C-terminus. Three disulfide bonds (C25-C79, C72-C96 and C27-C105) connect the core structural elements to the N-terminal region and stabilize the fold.

The $SYCY1_{RBD}$ forms an unexpected and extensive binding interface with the tranDs and scaDs of two $ASCT2_{WT}$ subunits (A and B, respectively) encompassing an area of nearly 3,000 $Å^2$, while forming only weak contacts with the scaD of subunit C (Fig. 1-c, Supplementary Information). The core and N-terminal regions of the $SYCY1_{RBD}$ wrap around the extracellular β-hairpin in the ASCT2 subunit A (Fig. 1a-c). Notably, the backbone chain of the $SYCY1_{RBD}$ segment F62-T66, establishes a complementary antiparallel β-strand with the extracellular β-hairpin of ASCT2 and forms a hybrid β-sheet with the receptor (Fig. 1c). We refer to this remarkable hybrid structural element as the "clamp", and to the N-terminal β-strand that complements the ASCT2 hairpin as the β_{CLAMP}.

The long α_C of the $SYCY1_{RBD}$ orients away from the ASCT2 vestibule and positions atop the tranD of subunit A, establishing numerous contacts with the transmembrane helix 6 (TM6), the helical-hairpin-2 (HP2), and TM8a (Fig. 1a-c). In turn, the short α_N of the $SYCY1_{RBD}$ packs against HP2 and TM8 in the tranD of the ASCT2 subunit B, while residues in the N-terminal (I47-R53) and core regions of the $SYCY1_{RBD}$ establish further contacts with the TM5 and the β-hairpin of the same ASCT2 subunit, respectively.

Overall, the $ASCT2_{WT}$-$SYCY1_{RBD}$ structure reveals that syncytin-1 recognizes its cellular receptor by burying its binding domain into the ASCT2 vestibule within the membrane plane, resembling the shape of a ball caught in a baseball glove, and stabilized by the strong backbone interactions at the clamp. Our results are in agreement with previous functional studies that established the N-terminal 124 residues of syncytin-1 as the minimal receptor binding domain[29], and demonstrate that the ASCT2 β-hairpin, rather than a negative determinant factor for syncytin-1 and RDR-retroviral recognition[39], is a key component of the binding interface.

Receptor-recognition mechanism of supressyn

The $SYCY1_{RBD}$ and $SUPYN_{RBD}$ share merely 20% amino acid sequence identity, and the latter lacks the $SDGGGX_2DX_2R$ signature motif, which was reported to be involved in receptor recognition[29] and is present in syncytin-1 (see below), as well as the envelopes of the RDR-group retroviruses (Extended Data Fig. 6a,b). Despite the evolutionary divergence, the structure of the $ASCT2_{CO}$-$SUPYN_{RBD}$ complex shows that the $SUPYN_{RBD}$ shares a conserved architecture with the $SYCY1_{RBD}$ (Fig. 2a). The core region of the $SUPYN_{RBD}$ superimposes accurately with that of the $SYCY1_{RBD}$ (root mean square deviation (rmsd) ~0.8Å), and the spatial distribution of the disulfide-bond network is preserved (Fig. 2b). Moreover, the core region of the $SUPYN_{RBD}$ orients in a similar way to that of the $SYCY1_{RBD}$ within the ASCT2 vestibule and likewise, the backbone of the $SUPYN_{RBD}$ N-terminal segment F61-T65 forms a

hybrid β-sheet clamp with the extracellular β-hairpin of the ASCT2 subunit A (Fig. 2c, d. Supplementary Information).

Beyond the above-mentioned similarities, the structure and binding mode of the $SUPYN_{RBD}$ differ in two important aspects from that of the $SYCY1_{RBD}$: first, the N-terminal region in the $SUPYN_{RBD}$ is significantly shorter and lacks the α_N, while its α_C orients nearly normal to the membrane plane. As a consequence, the $SUPYN_{RBD}$ forms only a few contacts with the extracellular loops of the tranD in the ASCT2 subunit A; second, the α_C of the $SUPYN_{RBD}$ packs extensively against the β-hairpin of ASCT2 in the subunit C, which bends towards the core region of the $SUPYN_{RBD}$, and the two proteins establish a short β-strand complementation between SUPYN-β1 and the ASCT2 β-hairpin (Fig. 2d). Hence, the $SUPYN_{RBD}$ is clamped at two sites through hybrid β-strand complementation with the ASCT2 β-hairpins of the subunits A and C, respectively, and only the former is structurally conserved with the $SYCY1_{RBD}$. The extended $SUPYN_{RBD}$ clamp system, compared to that of the $SYCY1_{RBD}$, likely contributes to its higher apparent affinity for ASCT2 (Extended Data Fig. 2c), and facilitates the negative regulation of syncytin-1 fusogenic activity (ref. 4 and Extended Data Fig. 1b).

The structures of the ASCT2 complexes with the $SYCY1_{RBD}$ and $SUPYN_{RBD}$ show similar architectural features of the two RBDs and overlapping recognition interfaces on the receptor that preclude simultaneous binding of the two retrovirus-derived proteins, revealing the structural mechanism that underlies the inhibitory effect of supressyn on the cell fusogenic function of syncytin-1[4].

Receptor-recognition mechanism of extant retroviruses

Retroviruses of the RDR superinfection interference group are able to restrict infections of members of the group through poorly understood receptor interference mechanisms[24], while supressyn blocks infections of several retroviruses of the RDR group that recognize the human ASCT2[10]. To shed light into the receptor-recognition mechanisms of the extant retroviruses, we

used artificial intelligence-based AlphaFold2 (AF)[46] to predict the structures of the RBDs of representative RDR-group retroviruses (RV_{RBD}), including the simian retrovirus 3 (SRV3, also known as Mason-Pfizer retrovirus) and the avian reticuloendotheliosis virus (REV) (Fig. 3a and Extended Data Fig. 6b). The accuracy of AF to predict a protein fold depends on several factors, and among the most important ones are the number of homologous amino acid sequences available, as well as the existence of experimental structures with similar folds. There is a significant number of RV_{RBD} homologous sequences from exogenous and endogenous retroviruses across several vertebrate species, but the PDB lacks experimental structures of those homologs. Therefore, as a control of the AF accuracy to predict the structures of the RV_{RBD}s, we compared the AF models of the $SYCY1_{RBD}$ and $SUPYN_{RBD}$ ($SYCY1_{RBD}$-AF and $SUPYN_{RBD}$-AF) to the corresponding cryo-EM structures determined in this work (Fig. 3a). Our AF models predicted well the general architecture of the RBDs and the structures of their individual domains and superimposed accurately at the level of the core and β_{CLAMP} regions (rmsd ~0.5Å) (Fig. 3a). They also agree well with previously reported AF models of the SYCY1 and RDR-retroviruses surface subunits[47]. However, there were significant differences in the relative orientations of the domains in both the $SYCY1_{RBD}$-AF (at the α_N) and $SUPYN_{RBD}$-AF (at the α_C) yielding overall rmsd values of 4.2Å and 5.4Å, respectively.

Notably, the AF structural predictions of the RV_{RBD}s (RV_{RBD}-AFs) reproduced accurately the experimentally determined fold of $SYCY1_{RBD}$ (Fig. 3a). All the RV_{RBD}-AFs show high degree of structural conservation at the core, the disulfide-bond network, and the β_{CLAMP} with both the $SYCY1_{RBD}$ and $SUPYN_{RBD}$ (rmsd ~1Å). The structural conservation, despite poor conservation of primary sequence (<25%), strongly argues that the RV_{RBD}s recognize the ASCT2 orthologs through the same molecular mechanism, likely establishing a similar 3D arrangement with the transporter within the vestibule, and stabilized by the formation of a hybrid β-sheet clamp, as revealed by the $ASCT2_{WT}$-$SYCY1_{RBD}$ cryo-EM structure.

Interestingly, all the highly conserved residues among the RV_{RBD} and $SYCY1_{RBD}$ map outside the receptor-binding interface in the $ASCT2_{WT}$-$SYCY1_{RBD}$ complex (Fig. 3b). Those signature residues include the cysteine pairs that form the disulfide bond network between the core and the N-terminal region, as well as the $^{115}SDGGGX_2DX_2R$ motif that connects the core to the α_C (Fig. 3c). The fact that the conserved residues among $SYCY1_{RBD}$ and the RV_{RBD} are not involved in receptor recognition raises two important questions: first, what the role of the conserved residues is. The triad of disulfide bonds and the $^{115}SDGGGX_2DX_2R$ motif likely confer stability to the fold and aid to position critical regions for contact with the receptor. Supporting this view, residues within the $^{115}SDGGGX_2DX_2R$ motif form a conserved and intricated hydrogen-bond network that links the core to the α_C (Fig. 3c,d), and likely determine its position, as well as dynamics. Second, how the vertebrate retroviruses of the RDR-group are able to recognize the human ASCT2 using evolutionary divergent RBDs. We argue that one of the key structural elements of the ASCT2 recognition mechanism is the hybrid β-sheet clamp formed between the N-terminal residues of the RBDs and the β-hairpin of ASCT2, and because this structure is mainly formed by backbone-atom interactions, it can tolerate certain degree of amino acid divergence. Therefore, the formation of a "clamp" as the one observed in the $ASTC2_{WT}$-$SYCY1_{RBD}$ complex is likely responsible for the cross-species recognition of human ASCT2 by the retroviruses of the RDR-group.

The structural conservation among the RBDs of endogenous and exogenous retroviruses is indicative of a common and ancient mechanism used for binding to ASCT2, whereby the RBDs bind overlapping surfaces on the ASCT2 vestibule and are clamped to the receptor through hybrid β-strand complementation. This mechanism likely contributes to the receptor interference phenomena observed among RDR-group retroviruses[8,24], as well as that between endogenous and exogenous retrovirus-derived proteins that use ASCT2 as the cellular receptor.

Allosteric modulation of ASCT2 transport

242 The transport mechanism of ASCT2 involves elevator-like motions of the tranDs across the
243 membrane to translocate the substrate and the sodium ions[37,38]. The tranDs of the three subunits
244 work independently, and binding of the substrate to the tranD is controlled by the movements
245 of HP2 that acts as a gate on opposite sides of the membrane[48,49].
246 The reported cryo-EM structure of $ASCT2_{WT}$ showed that in the absence of protein-binding
247 partners in detergent solutions, the tranDs of the three subunits adopt inward-facing states with
248 the substrate occluded within, and oriented towards the cytoplasm[33]. In contrast, in the
249 $ASCT2_{WT}$-$SYCY1_{RBD}$ complex, the tranDs of the subunits A and B have undergone large
250 movements towards the extracellular side and adopt outward-facing conformations (Fig. 4a).
251 Notably, these tranDs pack extensively against the $SYCY1_{RBD}$, while the tranD of the subunit
252 C, which is distant from the $SYCY1_{RBD}$, is in an inward-facing state. In turn, in the $ASCT2_{CO}$-
253 $SUPYN_{RBD}$ complex, the three tranDs are in inward-facing states. However, the tranD of the
254 subunit A that establishes limited contacts with the $SUPYN_{RBD}$ adopts an open position lacking
255 the bound substrate (Fig. 4b). Therefore, the ASCT2 structural changes occurring upon binding
256 of the retrovirus-derived proteins suggest that they could affect the dynamics of the tranDs, and
257 partially inhibit ASCT2 transport.
258 To test this hypothesis, we reconstituted purified $ASCT2_{CO}$ into proteoliposomes to probe the
259 effect of $SYCY1_{RBD}$ and $SUPYN_{RBD}$ on its sodium-dependent alanine transport function (Fig.
260 5a). Proteoliposomes loaded with 10 mM unlabeled alanine showed robust exchange with
261 external radiolabeled [^{3}H]-alanine in the presence of NaCl, whereas the absence of NaCl did
262 not support transport. The application of the $SYCY1_{RBD}$ and $SUPYN_{RBD}$ slightly decreased the
263 transport activity. However, to unmask the extent of their inhibitory effects, we had to overcome
264 the orientation problem of the reconstituted transporters. Using a fluorescence-based assay (Fig.
265 5b and Extended Data Fig. 7a), we determined that nearly 45% of the transporters were oriented
266 inside-out (with the cytoplasmic surface towards the extra-liposomal side), and while they

contributed significantly to the transport signal, they were unable to bind the $SYCY1_{RBD}$ or $SUPYN_{RBD}$.

To inhibit transport mediated exclusively by the inside-out transporters, we developed a nanobody, Nb469, that recognizes the cytoplasmic surface of ASCT2. Isothermal titration calorimetry experiments demonstrated that the Nb469 binds $ASCT2_{CO}$ with a dissociation constant (K_D) of 360 nM, and a 1:1 (Nb:ASCT2 protomer) stoichiometry (Extended Data Fig. 7b). The cryo-EM structure of the $ASCT2_{WT}$-Nb469 complex further revealed that the nanobody is wedged between the tranD and the scaD on the cytoplasmic side in each subunit, stabilizing an inward-facing state (Fig. 5c,d). The $ASCT2_{WT}$-Nb469 complex structure was solved at an overall 2.3Å resolution, the highest resolution reported for a mammalian SCL1 transporter (Extended Data Fig. 8 and Extended Data Fig. 9), enabling confident modelling of the sodium ions, structural water molecules, and bound substrate (L-Ala), and providing insights on the ASCT2 transport mechanism (see Supplementary Information).

Addition of Nb469 to the proteoliposomes inhibited approximately 34% of the total transport signal (Fig. 5a), which correlates well with the fraction of the inside-out oriented transporters. From the above results, we concluded that the Nb469 is an efficient transport inhibitor that selectively recognizes the cytoplasmic surface of ASCT2. Incubation of the Nb469-treated proteoliposomes with $SYNCY1_{RBD}$ or $SUPYN_{RBD}$ further decreased total transport, and inhibited the transport signal from "outside-out" transporters by approximately 25% and 17% (Fig. 5a), respectively. Consistently, electrophysiological measurements of the Na^+-induced substrate binding and the associated $ASCT2_{CO}$ conformational changes were decreased by approximately 43% and 28% in the presence of $SYNCY1_{RBD}$ or $SUPYN_{RBD}$, respectively (Fig. 5e). The functional results demonstrate that binding of the endogenous-retroviral RBDs to ASCT2 affect the tranD dynamics and slows down transport. Moreover, the inhibitory effects of the $SYCY1_{RBD}$ on both substrate uptake and transient currents were more pronounced than

292 those of the $SUPYN_{RBD}$, likely reflecting the fact that the $SYCY1_{RBD}$ interacts extensively with
293 and affects the dynamics of two tranDs (subunit A and B, burying 1,160 $Å^2$ surface area of the
294 domains), while $SUPYN_{RBD}$ weakly interacts with a single tranD (burying ~330 $Å^2$), as
295 observed in the cryo-EM structures.
296 Overall, the structural and functional results reveal that the HERV-derived proteins act as first-
297 in-class selective and partial inhibitors of ASCT2 transport, and suggest that the SYCY1 and
298 SUPYN expression in placenta downregulates ASCT2-mediated uptake of neutral amino acids.
299 **Discussion**
300 Our work reveals a common receptor-recognition mechanism by a diverse group of endogenous
301 and disease-causing animal retroviruses that use ASCT2 as cellular receptor. The mechanism
302 is at the core of cell-cell fusion events in placenta morphogenesis, the cell-entry and
303 superinfection-interference processes of the RDR-group animal retroviruses, as well as the
304 immunogenic function of SUPYN. There are four key structural features in the recognition
305 mechanism that have important functional implications. First, endogenous and exogenous
306 retroviral envelope proteins recognize the ASCT2 fold rather than a conserved amino acid
307 signature motif. ASCT2 exposes an unusually large area on the surface of the cell compared to
308 other vertebrate plasma membrane proteins and is able to re-shape its extracellular vestibule
309 through movements of the tranDs and β-hairpins. As observed in the $SYCY1_{RBD}$ and
310 $SUPYN_{RBD}$ complexes (Fig. 4a,b), the size and dynamics of the ASCT2 vestibule enable
311 recognition by endogenous and exogenous retroviral RBDs with variable structural elements
312 and amino acid sequences (Fig. 3a).
313 Second, binding of the individual RBDs is stabilized by the formation of a structurally-
314 conserved clamp. This hybrid β-sheet clamp is mainly made of backbone-atom interactions,
315 and can be formed by evolutionary divergent regions in both the RBD and ASCT2. Therefore,
316 the clamp mechanism is able to accommodate interspecies amino acid variations among ASCT2

orthologs, and facilitate cross-species infections among extant RDR-group retroviruses, as well as the acquirement of homologous Envs in taxonomically diverse retroviruses that use ASCT2 as the cellular receptor[24].

Third, the trimeric ASCT2 is able to bind only one RBD molecule in the vestibule, and competition for this binding site among different envelope proteins contributes to explain complex receptor-interference cellular phenomena underlying supressyn inhibition of cell-cell fusion[4], and its antiviral activity against the RDR-group retroviruses[10].

Fourth, the bound RBD locates deep within the membrane plane and is buried into the ASCT2 vestibule. In the context of the full-length fusogenic envelope trimers, this implies that at least one of the three RBDs should be fully solvent-exposed to engage with the receptor. Hence, it is possible that syncyntin-1 and the envelopes of the RDR-group retroviruses alternate between "open" and "close" pre-fusion conformations in which the RBD is exposed for receptor binding and protected within the envelope trimer, respectively, as it has been recently shown for the SARS-CoV-2 spike[50]. The only other reported structure of a gamma-type RBD bound to its receptor is that of the human syncytin-2 (HERV-FDR)[41], which is another functional membrane fusogen expressed in primate placentae[51]. Both the syncytin-2 RBD and its receptor, the lipid transporter MFSD2A[52], are structurally and evolutionary divergent compared to those of syncytin-1. Yet, the receptor-recognition mechanisms of syncytin-1 and -2 reveal an important common feature, their binding interfaces localize to a large extent within the membrane plane. This contrasts the binding mode of extant human viruses, including HIV[53], Ebola[54], and SARS-CoV-2[55] that recognize prominent surface-exposed domains of their cellular (co-) receptors. Considering that most gamma-type envelopes use solute carriers lacking extracellular domains as receptors[20], it is possible that the "intramembrane" receptor-binding interface is a universal feature among these envelope proteins. Moreover, this feature could constitute an evolutionary advantage to facilitate cell-cell fusion, over other types of retroviral envelope proteins – because

it would preclude "lateral" and futile binding events between the envelope and receptor molecules expressed on the surface of the same trophoblastic cell –, and help explain why many mammalian lineages co-opted gamma-type envelopes for the syncytin cell-fusogenic function. ASCT2 is upregulated in several forms of cancer and constitutes an emergent drug target in cancer therapy[31,32]. Despite current efforts to develop small-compound molecules [56], potent and selective transport inhibitors have not been reported. We show here that the $SYCY1_{RBD}$ and $SUPYN_{RBD}$ act as partial inhibitors of ASCT2 transport (Fig. 5a), and the structures we present provide a basis for optimization of minimal RBD constructs or design of RBD-based peptides with desirable potency, selectivity and efficacy through rational design. Moreover, the inhibitory effect of Nb469 and the accompanying structure show that nanobodies that block the tranD movements constitute an attractive alternative to develop ASCT2 inhibitors for cancer therapy.

Methods

ASCT2 expression and purification

Genes encoding ASCT2 proteins were synthesized (GenScript) and subcloned into pcDNA3.1(+) vector with N-terminal Strep-tag II affinity tag followed by a PreScission protease cleavage site. When GFP was required, it was cloned between the strep-tag and the PreScission site. Proteins were expressed in HEK-293F cells (Thermo Fischer) by transient transfection as previously described[36]. Cell pellets were resuspended in buffer containing 40 mM HEPES/Tris-base, pH 7.4, 200 mM NaCl, 1 mM EDTA, 5 mM L-Alanine, 1 mM PMSF and mammalian protease inhibitor cocktail (Sigma) and dounce-homogenized. Homogenized cells were then supplemented with 5% v/v glycerol, 1% dodecanoyl sucrose (DDS, Anatrace), 0.2% Cholesteryl Hemisuccinate Tris Salt (CHS, Anatrace) and incubated for 1 h with gentle stirring. Detergent-solubilized proteins were purified by affinity chromatography using StrepTactin Sepharose resin (GE Healthcare Life Sciences). Resin was pre-equilibrated and washed with buffer A containing 50 mM HEPES/NaOH pH 7.4, 5 mM L-Ala, 200 mM NaCl, 0.05 % DDS, and 0.01 CHS. The purified protein was eluted with buffer A supplemented with 2.5 mM D-desthiobiotin (Sigma). Peak fractions were concentrated in a 100-kDa molecular weight cut off (MWCO) Amicon Ultra filters (Merck Millipore) and injected on a size-exclusion chromatography (SEC) Superose 6 10/300 column (GE Healthcare), pre-equilibrated in SEC buffer containing 25 mM HEPES/Tris-base, pH 7.4, 100 mM NaCl, 5 mM L-alanine, 0.05% DDS, 0.01% CHS. The purified transporters were immediately used or flash-frozen and stored at -80 °C. All the purification steps were carried out at 4 °C. For reconstitution into liposomes and complexation with Nb469, the ASCT2 constructs were purified as described above, but all the affinity and SEC buffers were supplemented with 25 μM of a lipid mixture made of POPC, POPE and POPG (Avanti Polar Lipids) at 1:1:1 molar ratio.

HERV-derived protein purification

The codon-optimized human $SYCY1_{\Delta 439}$ and SUPYN genes were cloned into a modified pMTA/BiP vector (Invitrogen) for expression and secretion in *Drosophila melanogaster* Schneider 2 (S2) cells (Invitrogen) with a C-terminal PreScission protease cleavage site followed by a C-tag affinity tag (Thermofisher Scientific). The sequence of $SYCY1_{\Delta 439}$ included two mutations that are outside the RBD region observed in the structure of the $ASCT2_{WT}$-$SYCY1_{RBD}$ complex. Mutation C186S was introduced to avoid the detachment of the surface and transmembrane subunits, while S307N was a consensus exchange from an alignment of SYCY1 hominoid orthologs to improve stability.

Drosophila melanogaster Schneider 2 (S2) cells were grown at 23°C under normal atmosphere in Schneider's Drosophila Medium (Thermofisher Scientific) supplemented with 10% of heat-inactivated Fetal Bovine Serum (FBS, Thermofisher). Cell density and viability were measured routinely in a TC20 Automated Cell Counter (Bio-Rad). $SYCY1_{\Delta 439}$ and SUPYN stable cell lines were generated using a puromycin-based selection method as described before[57,58] with modifications. Briefly, S2 cells were seeded at 1×10^6 cells/ml and co-transfected with the gene of interest and pCoPURO plasmid using Effectene (Qiagen). Puromycin selection (10 μg/mL) was done for over 20 days and then, cells were adapted to the Insect-Xpress serum-free media supplemented with L-glutamine (Lonza), and transferred to an orbital shaker to grow in suspension at 23°C.

Suspension-cell cultures at ~$5x10^6$ cells/ml were induced using 4 μM $CdCl_2$ for 4 days at 23°C. Cells were pelleted at 4°C and the supernatant was filtered through a 0.22 μm filter and mixed with the CaptureSelect C-tagXL resin (Thermo Scientific) equilibrated with buffer containing: 50 mM HEPES pH 7.4/Tris-base, 200 mM NaCl, 5 mM alanine and 5% glycerol (buffer A). The purified proteins were eluted with buffer A supplemented with 2 mM of the S-E-P-E-A peptide (Genscript), and concentrated using a 3-kDa MWCO Amicon Ultra filters (Merck Millipore) for SUPYN, and 30-kDa MWCO concentrator $SYCY1_{\Delta 439}$. The concentrated

proteins were further purified by SEC using Superdex 75 Increase 10/300 (Cytiva) for SUPYN and Superose 6 Increase 10/300 columns (Cytiva) for SYCY1$_{\Delta 439}$ equilibrated with buffer A. All purification steps were carried out at 4°C. For complexation with purified ASCT2 proteins the SEC-purified SUPYN and SYCY1$_{\Delta 439}$ were supplemented with 0.05 % DDS.

Nanobody generation, expression, and purification

ASCT2 specific nanobodies were generated as previously described[59]. For the immunization, ASCT2 carrying an N-terminal Strep-tag II affinity tag followed by GFP and a PreScission protease cleavage site were cleaved at the PreScission site to eliminate the GFP and the affinity tag. In brief, a llama (*Lama glama*) was immunized six times with a total of 0,8 mg of ASCT2$_{WT}$. Four days after the final boost, blood was taken from each llama to isolate peripheral blood lymphocytes. RNA was purified from these lymphocytes and reverse transcribed by PCR to obtain the cDNA of the ORFs coding for the nanobodies. The resulting library was amplified by PCR and cloned into the pMESy4 phage display vector bearing a C-terminal hexa-His tag and a CaptureSelect C-tag sequence (Glu-Pro-Glu-Ala). Both proteins were solid phase coated directly on Maxisorp plates in 50mM HEPES pH7,4, 200mM NaCl, in the presence of 10mM Gln and selections were done in the same buffer. ASCT2$_{WT}$ specific phage were recovered by limited trypsinization. After two rounds of selection, periplasmic extracts were made and subjected to ELISA screens in the same conditions as described above. The Nb469 was selected among different Nb families based on a qualitative binding assay in which the Nbs were immobilized by the CaptureSelect C-tag affinity tag, and the fluorescence of the bound GFP-ASCT2 fusion construct was taken as a proxy of the Nb binding potencies.

Protein binding assays

To probe protein-protein binding, we used the GFP-ASCT2$_{CO}$ construct because of the higher expression yields and protein stability compared to the GFP-ASCT2$_{WT}$. Purified SUPYN and SYCY1$_{\Delta 439}$ were supplemented with DDS, and immobilized onto the CaptureSelect affinity

resin equilibrated with buffer: 50 mM HEPES/NaOH pH 7.4, 5 mM L-Ala, 200 mM NaCl, 0.05 % DDS, and 0.01 CHS. Purified GFP-ASCT2$_{CO}$ at different concentrations was incubated with the resin for 30 min at 4°C while being mixed by rotation. The resin was washed with 8 column-volumes of the above-mentioned buffer and then eluted by supplementing the buffer with 2 mM of the S-E-P-E-A peptide. The eluted samples were transferred into a 96-well plate for quantification of GFP fluorescence (470nm/514nm, $\lambda_{excitation}/\lambda_{emission}$) using a CLARIOstar-Plus microplate reader (BMG Labtech). As controls for estimation of the background fluorescence, we used affinity resin without HERV-derived protein bound and treated with GFP-ASCT2$_{CO}$. After subtracting the background fluorescence, the data were fitted to a Hill equation of the form:

$$\text{(SYCY1 or SUPYN) Fraction bound} = [\text{ASCT2}]^h/(K_D{}^h + [\text{ASCT2}]^h)$$

Where K_D is the apparent dissociation binding constant, and h is the hill slope.

Binding of the Nb469 to GFP-ASCT2$_{CO}$ was measured by isothermal titration calorimetry using a MicroCal iTC200 calorimeter (Malvern Panalytical). The two proteins were purified with a SEC buffer containing: 20 mM HEPES, 100 mM NaCl, 5 mM L-Ala, 0.05% DDS, and 0.01% CHS. The Nb469 was loaded in the syringe at a concentration ~7-fold higher than that of the transporter in the experimental cell. The experiments were carried out at 25°C. Protein concentration was determined by absorbance at 280 nm. The data were fitted to the quadratic binding equation ("one site binding model") using built-in routines in Origin 7 (OriginLab).

Cell fusion assay

Cell-cell fusion was assayed as described before[41] using split-GFP complementation[60]. Briefly, the GFP β-strand 11 (GFP11) was fused to the C terminus of the ASCT2 constructs, and the gene encoding the GFP β-strands 1–10 (GFP1–10) was cloned into pcDNA3.1(+) vector. The proteins were expressed in HEK-293T.17 (ATCC) adherent cells that were grown in DMEM medium supplemented with 10% FBS (Gibco) at 37°C in 5% CO_2. Cell cultures were

transfected in 24-well plates using Lipofectamine 2000 (Invitrogen) according the manufacturer's instructions. Twenty-four hours after transfections, cells expressing SYNC1-GFP_{1-10} were gently resuspended and layered over attached cells expressing ASCT2-GFP_{11}. After 6-hour incubation at 37°C in 5% CO_2, the cells were detached, washed with PBS, and transferred to a 96-well plate for fluorescence quantification. GFP fluorescence was measured in a CLARIOstar-Plus plate reader (BMG Labtech) (470nm/514nm, $\lambda_{excitation}/\lambda_{emission}$). To probe the effect of the HERV-derived proteins, the cells expressing the ASCT2-GFP_{11} fusions were pre-incubated with 10µM purified SUPYN or $SYCY1_{\Delta 439}$ for 15 min at 37°C before assaying fusion.

Proteoliposomes transport and electrophysiology assays

The ASCT2 proteins were reconstituted in proteoliposomes in the presence of 50mM HEPES pH 7.4, 200 mM NaCl, 10mM Gln. Proteoliposomes were resuspended in a buffer containing 40 mM HEPES/Tris-base pH 7.4, 100 mM NaCl and 10 mM L-Ala. The resuspended samples were subjected to freeze/thaw cycles followed by extrusion using 0.4 µm filters. The sample was ultracentrifuged and resuspended in buffer containing 40 mM HEPES/Tris-base pH 7.4 and 100 mM Choline-Cl (buffer W). The transport reaction was initiated by adding 4 µL of the resuspended proteoliposomes to 40 µL of 40 mM HEPES/Tris-base pH 7.4, 100 mM NaCl, 50 uM L-Ala, and 2 µL of 33 µM [2,3-^{3}H]-L-Ala (Hartmann Analytic) (buffer B). The tubes were incubated for 30 minutes at 25ºC with mild shaking. The reaction was diluted by adding 160 µL of buffer W. A negative control using buffer B, where NaCl was replaced by Choline-Cl, was performed in an identical manner. The samples were rapidly filtered using MultiScreenHTS filter plates (Merk), followed by washing with buffer W. Radioactivity levels were measured by Liquid Scintillation counting using an Analyzer Tri-Carb 4910 TR counter (PerkinElmer). Inhibition assays were performed using 20 µM of SUPYN and $SYCY1_{\Delta 439}$, and

12 µM of Nb469. In these assays, proteoliposomes were incubated for 15 to 30 minutes with the protein binders before starting the transport reaction.

For determination of GFP-ASCT2$_{CO}$ orientation in the proteoliposomes membrane, samples were first treated as mentioned above in absence of radiolabelled substrate. The proteoliposomes samples were incubated with an excess of 3C protease to cleave the GFP and incubated for about 16 hours at 10ºC. The reaction was terminated by resuspension in Laemmli buffer and proteins were separated by SDS-PAGE analysis. In gel GFP Fluorescence was measured using a Fusion FX system (Vilber Lourmat) (488/535 nm, $\lambda_{excitation}/\lambda_{emission}$). ImageJ v2.0.0 was used for analysis of fluorescence intensity of the gel bands.

Solid-Supported Membrane (SSM)-based electrophysiology was performed with a SURFE^{2}R N1 instrument (Nanion Technologies). Alkylation of 3-mm SURFE^{2}R N1 single sensor was done by incubation of 100 µl of a 0.5 mM thiol solution (1-octadecanethiol in isopropanol) for 1 hour at room temperature, followed by washes with isopropanol and Milli-Q water. The membrane was formed by applying 1.5µl of a lipid solution (7.5 µg/µl 1,2-diphytanoyl-sn-glycero-3-phosphocholine in n-decane) on the surface of the alkylated gold electrode followed by addition of 50 µl of a SSM buffer containing: 40 mM HEPES (pH 7.4), 100 mM Choline-Cl, and 5 mM L-alanine. The proteoliposome suspension (at a lipid concentration of 20 mg/ml) was diluted 1:10 in SSM buffer, and sonicated for 30 seconds. 10µl of the sonicated proteoliposomes was added to the sensor before centrifugation at 3000g for 30 min. To ensure the quality of the SSM, the conductance and capacitance were measured before every measurement and values lower than 5 nS and between 15-35 nF, respectively, were considered acceptable. The capacitive transients were elicited by exchanging the SSM buffer by the activating buffer containing: 50 µM L-alanine, 10 mM NaCl, 40 mM HEPES (pH 7.4), and 90 mM Choline-cl. SYCY1$_{\Delta 439}$ and SUPYN were preincubated for 1 hour with the proteoliposomes at room temperature.

Electron microscopy sample preparation and data acquisition

The ASCT2 complexes were formed at a ~1.1-molar excess of the $SYCY1_{\Delta 439}$ and SUPYN, and at ~1.2 molar excess of Nb469 over the ASCT2 monomeric concentration. The complexes were further purified by SEC using the above-mentioned buffers and concentrated to 4-9 mg/ml. 3-4 µl of complex samples were spotted on glow discharged Au 300 mesh Quantifoil R1.2/1.3 holey carbon grids (Quantifoil). Grids were blotted for 4 s at 4°C and 100% humidity and plunge-frozen into liquid ethane using a FEI Mark IV Vitrobot (Thermo Fisher). Grids were screened on a 200 kV Talos Arctica microscope (ThermoFisher) at the IECB cryo-EM imaging facility.

Data collections for the $ASCT2_{WT}$-$SYCY1_{RBD}$ and $ASTC2_{WT}$-Nb469 complexes were performed on a 300 kV Titan Krios G4 microscope (ThermoFisher) at the EMBL-Heidelberg Cryo-Electron Microscopy Service Platform, equipped with a Falcon 4i detector (ThermoFisher) operating in electron event representation mode (EER)[61], and using SerialEM software[62]. For the $ASCT2_{WT}$-$SYCY1_{RBD}$ complex, the acquisition was done with a pixel size of 0.731 Å/pixel (165,000 x magnification) and a total dose equivalent to 53.4 $e^-/Å^2$ in 1,512 EER frames in 6s, and with a target defocus range of -0.4 to -1.50 µm. For the $ASTC2_{WT}$-Nb469, the acquisition was done with a pixel size of 0.934 Å/pixel (130,000 x magnification) and a total dose equivalent to 52.24 $e^-/Å^2$ in 1,260 EER frames in 4s, and with a target defocus range of -0.6 to -1.50 µm.

The $ASCT2_{CO}$-$SUPYN_{RBD}$ data set was recorded on a Titan Krios electron microscope equipped with K3 direct electron detector (Gatan) operating in counting mode, and using SerialEM[62]. The pixel size was 0.814 Å with a target defocus range of -0.8 to -2.0 µm. Each movie contained 40 frames with a dose per frame of 1.03 $e^-/Å^2$, and a total exposure time of 8s.

Cryo-EM data processing, model building, and structure analysis

528 The datasets were processed with cryoSPARC[63] v3 ($ASCT2_{WT}$-$SYCY1_{RBD}$ and $ASCT2_{WT}$-
529 Nb369) or, v3 and v4 ($ASCT2_{CO}$-$SUPYN_{RBD}$) (Extended Data Table 1). Movies were gain
530 corrected, and aligned using patch motion correction. Contrast transfer function (CTF)
531 parameters were estimated using the patch-CTF routine within cryoSPARC.
532 For the $ASCT2_{WT}$-$SYCY1_{RBD}$ dataset (Extended Data Fig. 3), particles were initially picked
533 from a subset of 2,000 movies using the "blob picker" job and then 2D classified. The 2D
534 classes that appear to contain transporters within detergent micelles were then used as
535 references to pick particle across the entire dataset (14,000 movies). Over four million particles
536 were reference-picked, 2-fold binned, and extracted. Rounds of 2D classification yielded
537 640,000 particles that were re-extracted using the original pixel size. Iterative 3D "ab-initio"
538 classifications followed by non-uniform[64] and local refinements yielded an overall ~2.6Å
539 resolution map, based on the 'gold-standard' 0.143 FSC cut-off. The map was of excellent
540 quality over the $SYCY1_{RBD}$, and the majority of the transporter, but the tranD of the subunit C
541 that localizes far from the $SYCY1_{RBD}$ binding interface was poorly resolved. To overcome this
542 problem, we re-classified the 640,000 particles obtained from the 2D classifications using
543 heterogeneous refinement. Iteratively, we used maps with improved density for the third tranD
544 as the "good reference" against maps without molecular features ("bad class"), and obtained a
545 set of ~26,000 particles that after homogenous and local refinements yielded an overall 3.5Å-
546 resolution map. The map was sharpened using the DeepEMhancer[65] software and enabled to
547 confidently track the backbone of the third tranD.
548 To build the structure of the $ASCT2_{WT}$-$SYCY1_{RBD}$ complex, we first built the atomic model of
549 the transporter using the reported ASCT2 cryo-EM structures (PDB 6GCT[33] and 6MPB[34]) to
550 fit the tranDs and scaDs as rigid bodies into the density. Then the ASCT2 atomic model was
551 corrected manually using Coot[66] v0.9.8.3. The structure of the bound $SYCY1_{RBD}$ was manually
552 built *de novo*. The $ASCT2_{WT}$-$SYCY1_{RBD}$ structural model at ~2.6Å-resolution includes

553 syncytin-1 residues 21-141 encompassing its RBD, as well as most residues of the three ASCT2
554 scaDs and the tranDs in subunits A and B, except for the distal N- and C-termini, and a few
555 residues on the extracellular loop between TM3-TM4. To build the third tranD (subunit C), we
556 used the ~2.6Å-resolution structure as the starting model.
557 For the $ASCT2_{CO}$-$SUPYN_{RBD}$ dataset (Extended Data Fig. 4), ~1.5 million particles were
558 reference picked, 2-fold binned, and 2D-classified. Approximately, 400,000 selected particles
559 were used for ab-initio 3D classification. A good-quality map displaying secondary structural
560 features of both ASCT2 and $SUPYN_{RBD}$ was then used to re-process the entire set of picked
561 particles (1.5 M) using heterogenous refinement iteratively against a map without molecular
562 features. Then, ~125,000 selected particles were re-extracted to the original pixel size followed
563 by non-uniform refinement yielding a map at an overall ~3.6Å resolution. The refined particles
564 were used for 3D Flexible refinement that improved the map quality and the resolution to
565 ~3.4Å.
566 The model building of the $ASCT2_{CO}$-$SUPYN_{RBD}$ complex was done similarly to the $ASCT2_{WT}$-
567 $SYCY1_{RBD}$ complex. The $SUPYN_{RBD}$ model was manually built *de novo*, except for the N-
568 terminal extreme residues that we modeled with the help of the AlphaFold prediction
569 (M5A8F1)[67].
570 For the $ASCT2_{WT}$-Nb469 dataset (Extended Data Fig. 8), initial particles from 500 images were
571 "blob picked" and were used for 2D classification (2-fold binning). The selected 64,683
572 particles were used to train the in-built Topaz[68] job in cryoSPARC. Nearly, 1.3M particles were
573 picked with topaz over the entire data set (4,344 images), extracted after 2-fold binning and
574 subjected to 2D and 3D ab-initio classification. Selected ~500,000 particles were re-extracted
575 at the original pixel size, ab initio re-classified. A final set of ~450,000 particles were subjected
576 to non-uniform refinement and yielded a map at an overall 2.3Å resolution after imposing C3
577 symmetry.

578 Models of ASCT2 and a non-related nanobody (6GCT and 7PQQ PDB codes, respectively)
579 were superimposed in the final map of the complex and then modified in Coot[66]. Water
580 molecules were first located using the douse tool in Phenix v1.20.1[69] and then validated in Coot.
581 Final atomic coordinates of the complex were refined using Phenix v1.20.1 using secondary
582 structure and Ramachandran restraints and validated using the MolProbity Phenix validation
583 tool. Maps were visualized using UCSF Chimera[70].
584 The AlphaFold structural predictions were computed using the ColabFold server[71] with
585 default settings. The RBD amino acid sequences were retrieve from the following uniport
586 entries Q77YG6 (SRV3), A7LKA7 (RD114), P03399 (REV), and P31796 (SNV).
587

Acknowledgments

IECB cryo-EM imaging facility is acknowledged for support in cryo-EM sample screening and initial data acquisition, and the EMBL-Heidelberg Cryo-Electron Microscopy Service Platform for support in image acquisition. Reda Assal for preliminary protein production and construct screening. This research was funded by the European Research Council (ERC) under the European Union Horizon 2020 Program (Grant 771965), with additional contributions from the Institut National du Cancer (INCA grant 2017-44), IDEX Senior Chair Universite de Bordeaux, and Region Nouvelle-Aquitaine (Grant 8166910), all to NR. Instruct-ERIC (PID6507), as part of the European Strategy Forum on Research Infrastructures (ESFRI), and the Research Foundation-Flanders (FWO) for Nanobody discovery. Alison Lundqvist for technical assistance during Nanobody discovery. The Swiss National Science Foundation (SNSF) (PP00P3_198903 and 310030_207974 to CP) supported the work performed by the group of CP. FAR and MB acknowledge funding from grant ANR-10-LABX-62-IBEID.

Author contributions

SK, JC-T, AK performed protein purification, cryo-EM sample preparation and image analysis, as well as atomic model building and refinement. SK and MV performed binding assays, and MV performed cell fusion assays, as well as optimized, screened, and produced the HERV-derived proteins. MV and JC-T produced the proteoliposomes. ABC and CP performed and analyzed transport and electrophysiology assays with proteoliposomes and binders. MB and FR contributed to the design and development of soluble HERV constructs. EP and JS designed and performed nanobody development. NR analyzed functional and cryo-EM data, as well as structures. All authors contributed to the preparation of the manuscript. N.R conceived and supervised the project.

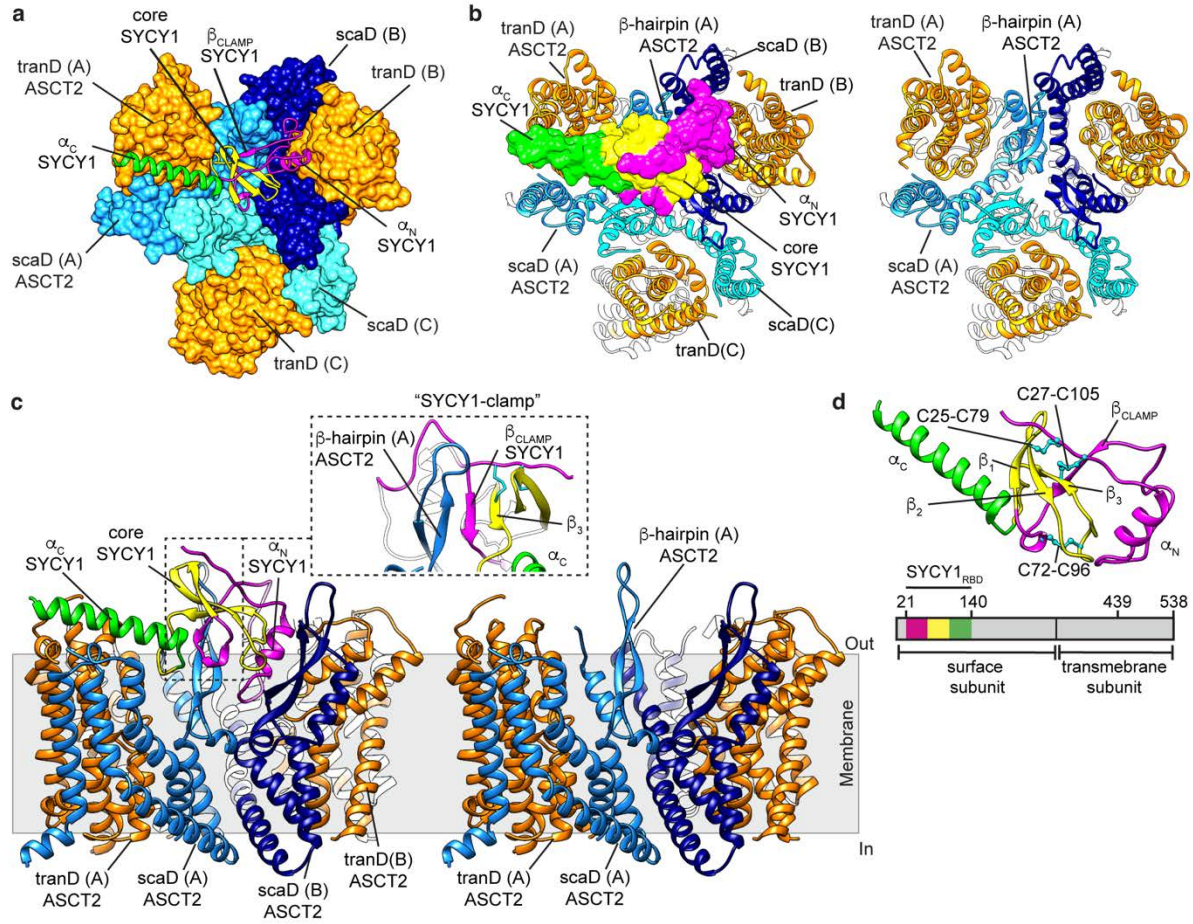
Competing interests

AK is currently an employee of the Integrated Drug Discovery, Sanofi R&D, France. All other authors declare no competing interests.

Data availability: Structural models of $ASCT2_{WT}$-$SYCY1_{RBD}$ at 2.62 Å and 3.5 Å resolution, $ASCT2_{CO}$-$SUPYN_{RBD}$, and $ASCT2_{WT}$-Nb469 complexes have been deposited to the PDB with accession codes 8OUH, 8OUJ, 8OUI and 8OUD, respectively, and the corresponding cryo-EM maps were deposited in the Electron Microscopy Data Bank (EMDB) under accession numbers EMD-17192, EMD-17194, EMD-17193 and EMD-17189. Materials are available upon reasonable request.

Corresponding author

Correspondence and requests for materials should be addressed to C.P. (camilo.perez@unibas.ch) or N.R. (nicolas.reyes@u-bordeaux.fr)



624

625 **Figure 1: Cryo-EM structure of ASCT2$_{WT}$-SYCY1$_{RBD}$ complex**
626 **a**, Extracellular view of the ASCT2$_{WT}$ (surface representation) in complex with the SYCY1$_{RBD}$
627 (ribbon representation). The three ASCT2$_{WT}$ tranDs are in orange and the central "propeller-
628 like" structure formed by the three scaDs in different tones of blue. The corresponding ASCT2
629 subunits are indicated in brackets. The SYCY1$_{RBD}$ N-terminal region is in pink, the core in
630 yellow, and the α_C in green. **b**, Extracellular view of the ASCT2$_{WT}$ (ribbon representation) in
631 complex with the SYCY1$_{RBD}$ (surface representation) (left). The structure of ASCT2$_{WT}$ in the
632 complex, after removing the SYCY1$_{RBD}$, is also shown for clarity (right). **c**, Membrane view
633 of the complex (left). The ASCT2$_{WT}$ subunit C and some elements of the TM4 in the scaD in
634 the subunit B were removed to display the binding interface. The structure of ASCT2$_{WT}$ in the
635 complex, after removing the SYCY1$_{RBD}$, is also shown for clarity (right). The inset shows the
636 details of the hybrid β-sheet clamp formed between the N-terminal β_{CLAMP} of the SYCY1$_{RBD}$
637 (pink) and the extracellular β-hairpin of ASCT2$_{WT}$ subunit A. **d**, Structure of the bound
638 SYCY1$_{RBD}$ highlighting the different structural elements with the same color code as in (b). A
639 linear diagram of the syncytin-1 architecture is shown.
640

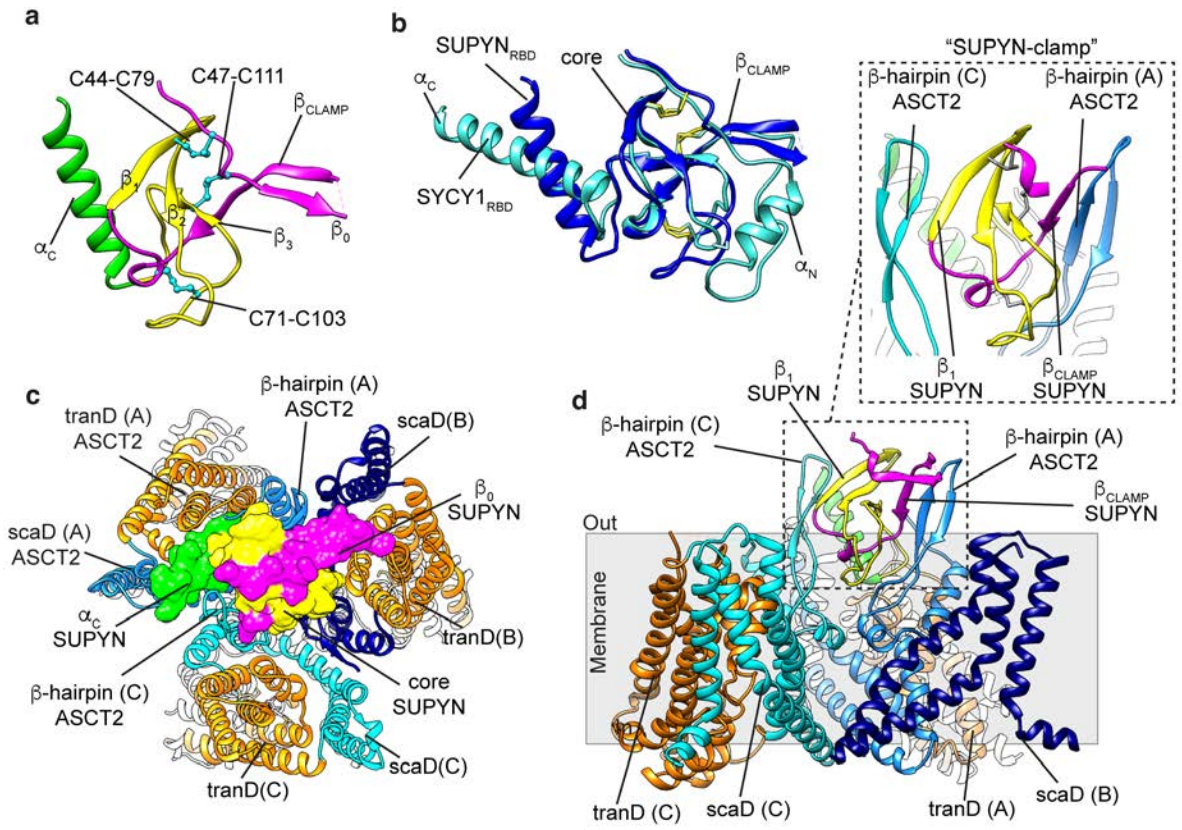


Figure 2: Cryo-EM structure of $ASCT2_{CO}$-$SUPYN_{RBD}$ complex

a, Structure of the bound $SUPYN_{RBD}$ is shown with the N-terminal region in pink, the core in yellow and the α_C in green. **b**, Superimposition of the bound $SYCY1_{RBD}$ (cyan) and $SUPYN_{RBD}$ (blue). The disulfide-bond network is represented as yellow sticks. **c**, Extracellular view of the $ASCT2_{CO}$ (ribbon representation) in complex with the $SUPYN_{RBD}$ (surface representation). The $ASCT2_{CO}$ tranDs are colored in orange and the scaDs of the three subunits in different tones of blue. The corresponding subunit of each domain is indicated in brackets. The different structural elements of the $SYCY1_{RBD}$ are color coded as in (a). **d**, Membrane view of the complex. The tranD of subunit B and some elements of its scaD were removed for clarity of display. The inset shows details of the $SUPYN_{RBD}$ clamp with two regions of hybrid β-strand complementation, one between SUPYN-β_{CLAMP} and the ASCT2 β-hairpin in the subunit A, and another between SUPYN-β_1 and the ASCT2 β-hairpin in the subunit C.

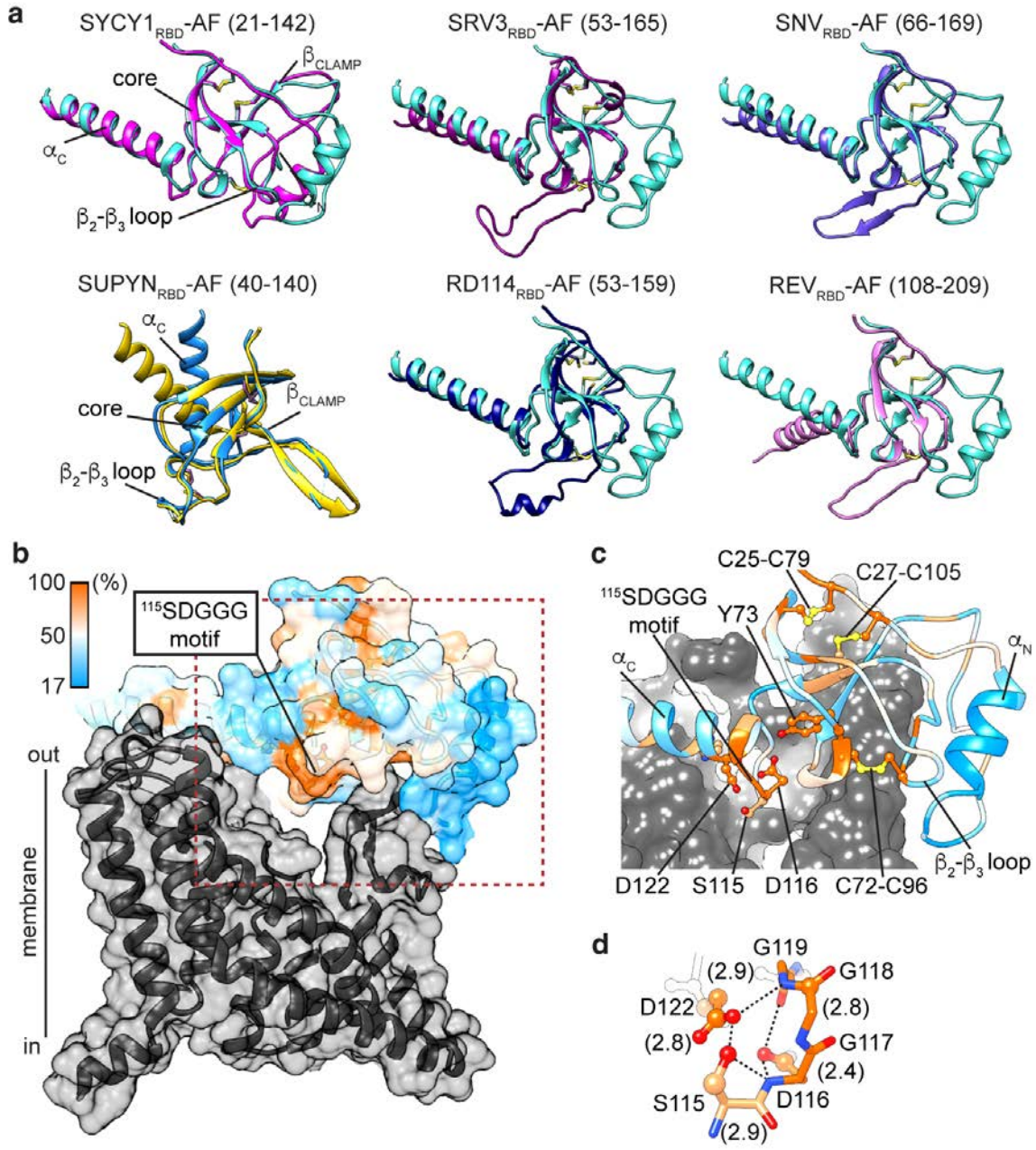


Figure 3: Structural comparison of endogenous and exogenous retroviral RBDs

a, The cryo-EM structure of the $ASCT2_{WT}$-bound $SYCY1_{RBD}$ (cyan) is superimposed to the AlphaFold model of the following RBDs: SYCY1 (pink), SRV3 (dark pink), SNV (purple), RD114 (blue), and REV (light pink). The superimposition of the $ASCT2_{CO}$-bound $SUPYN_{RBD}$ cryo-EM structure (light blue) to the $SUPYN_{RBD}$ AlphaFold model (yellow) is also shown. **b**, The amino acid conservation across the RV_{RBD}s and the $SYCY1_{RBD}$ is mapped on the cryo-EM structure of the bound $SYCY1_{RBD}$ to $ASCT2_{WT}$ (only subunit A is represented for clarity). The scale bar represents the percentage of amino acid conservation. **c**, The side chains of the most conserved residues are displayed and lack contacts with the receptor. **d**, Hydrogen-bond network among residues of $^{115}SDGGGX_2D$ motif (dashed lines). Side-chain and main-chain atoms are displayed as spheres and sticks, respectively. Interatomic distances in angstrom are indicated in parentheses.

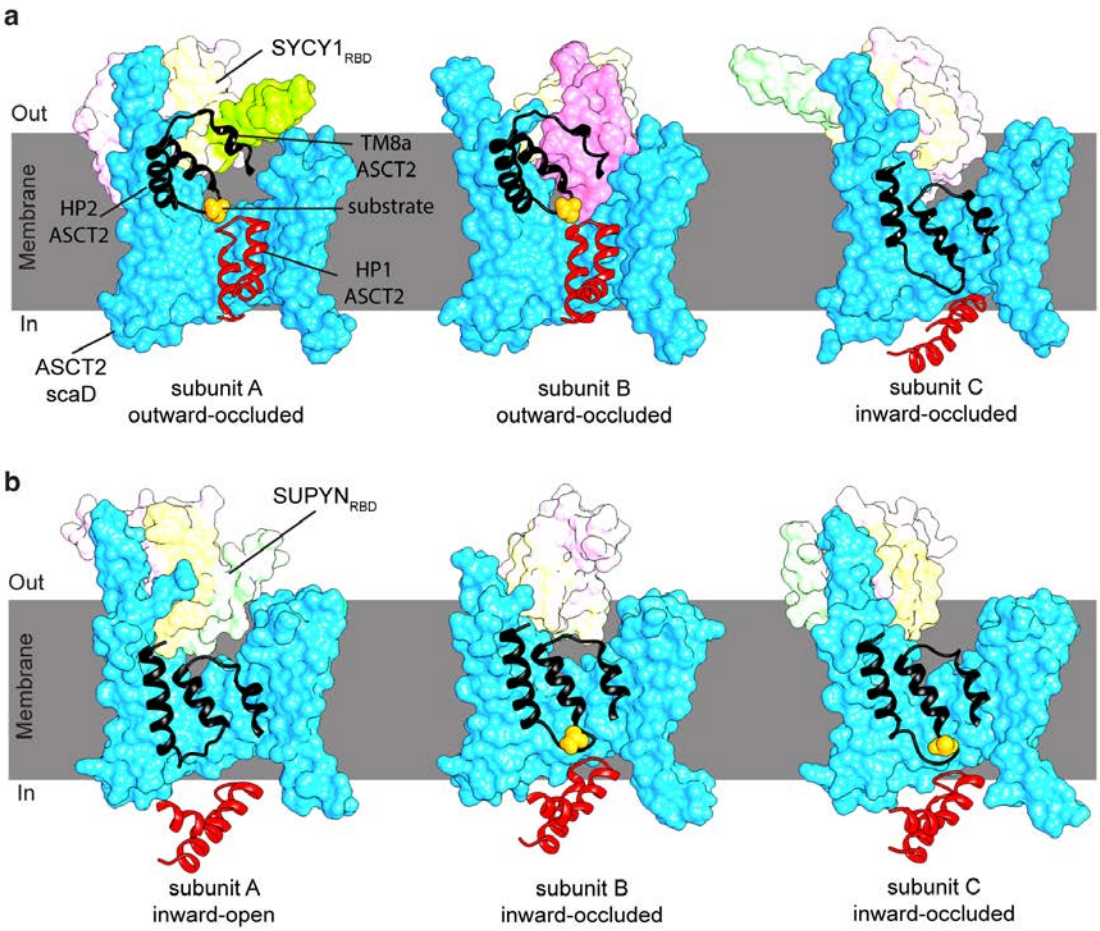


Figure 4: Conformational changes of the tranDs in the RBD-bound ASCT2 complexes
a, Membrane views of the individual $ASCT2_{WT}$ subunits with bound $SYCY1_{RBD}$. The subunits are rotated by 120° respect to each other around the pseudo-symmetry axis of the three scaDs. The scaDs (cyan) and the $SYCY1_{RBD}$ (N-terminal, pink; core, yellow; α_C, green) are shown as protein surfaces. The HP2-TM8 (black) and the HP1 (red) in the tranDs are shown as ribbons, and the L-ala substrate (yellow) is represented as spheres. The $SYCY1_{RBD}$ regions that interact with the tranDs are shown in a brighter color. **b**, Membrane views of the three $ASCT2_{CO}$ individual subunits in the $SUPYN_{RBD}$ complex. Color codes and representations are as in (a).

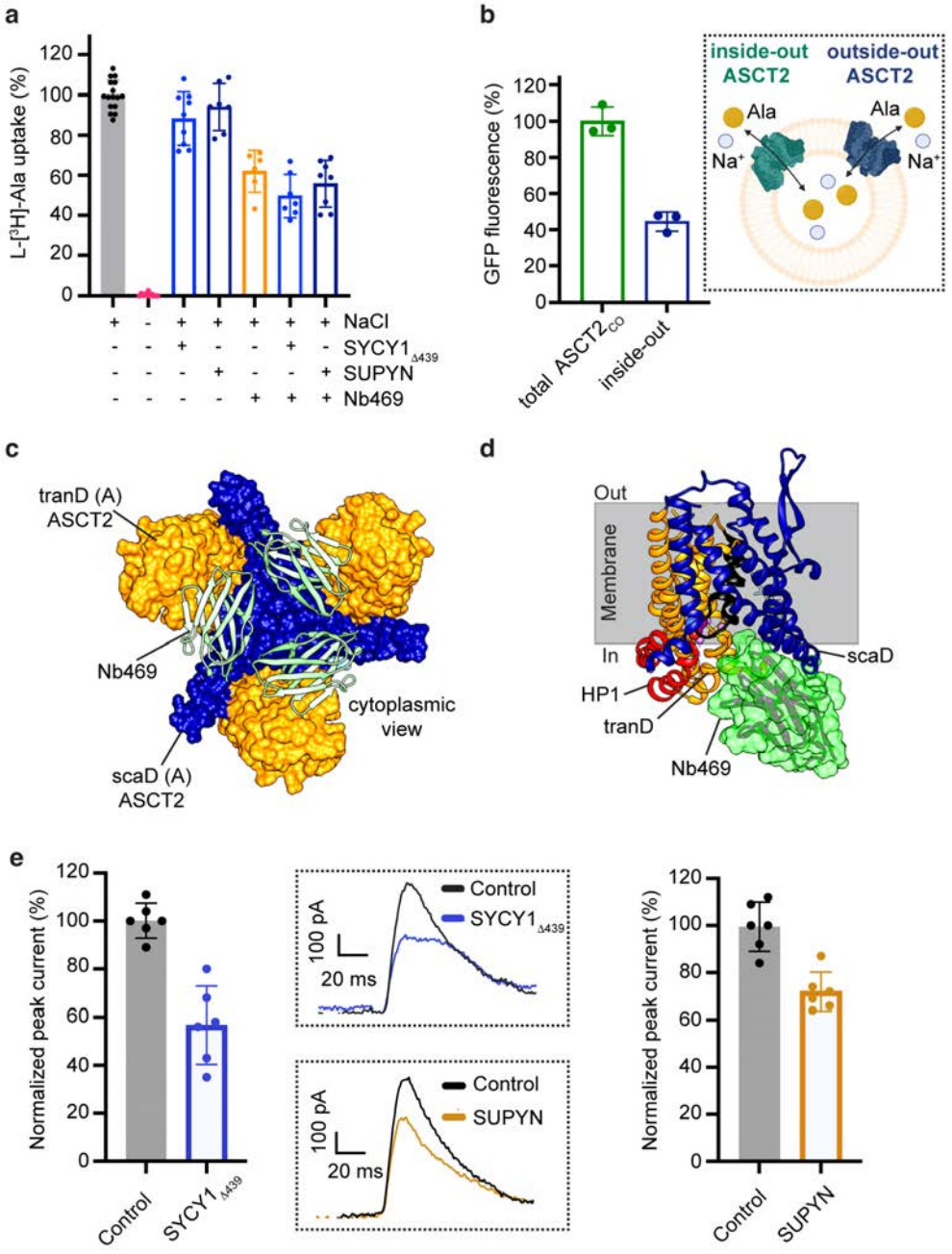
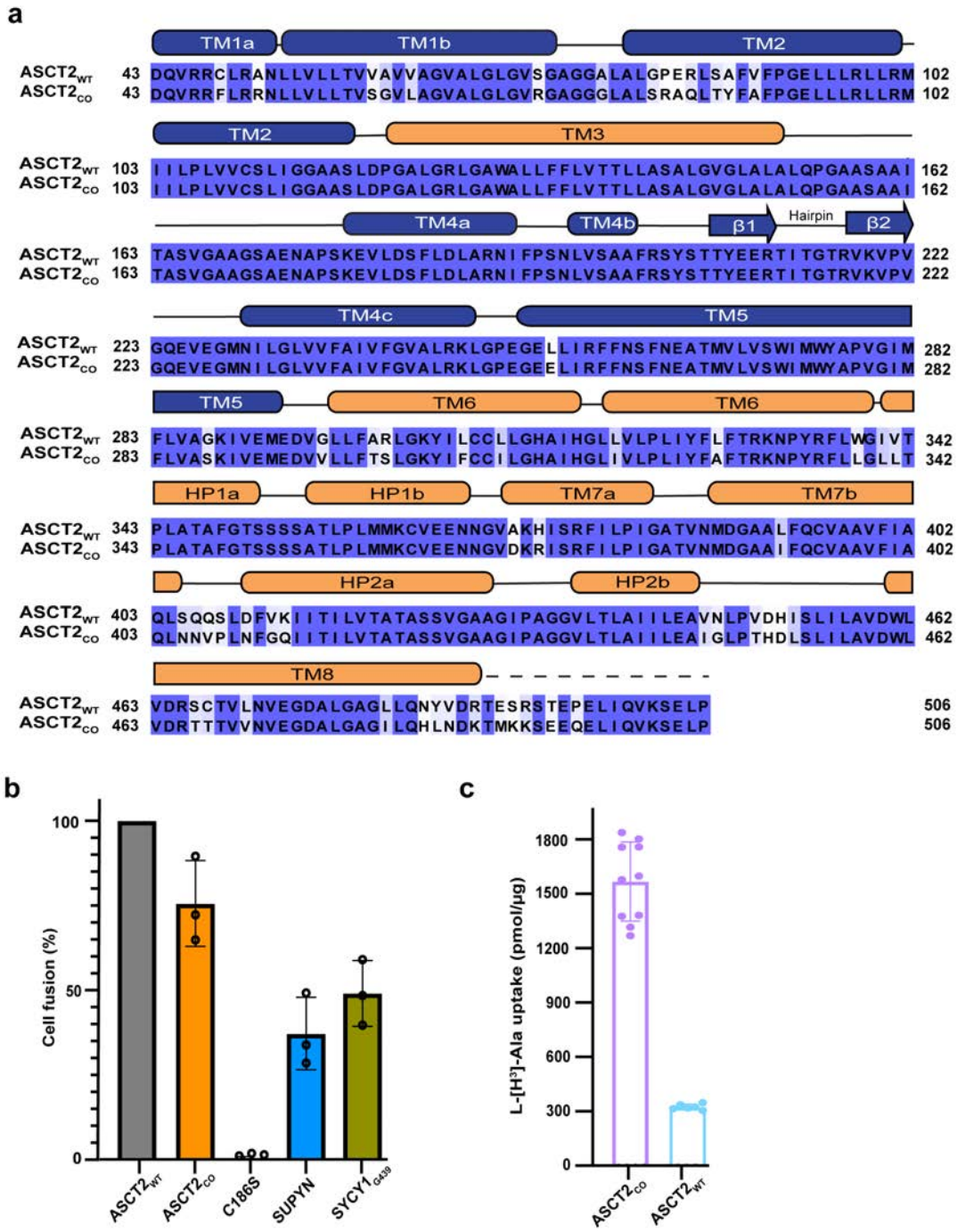


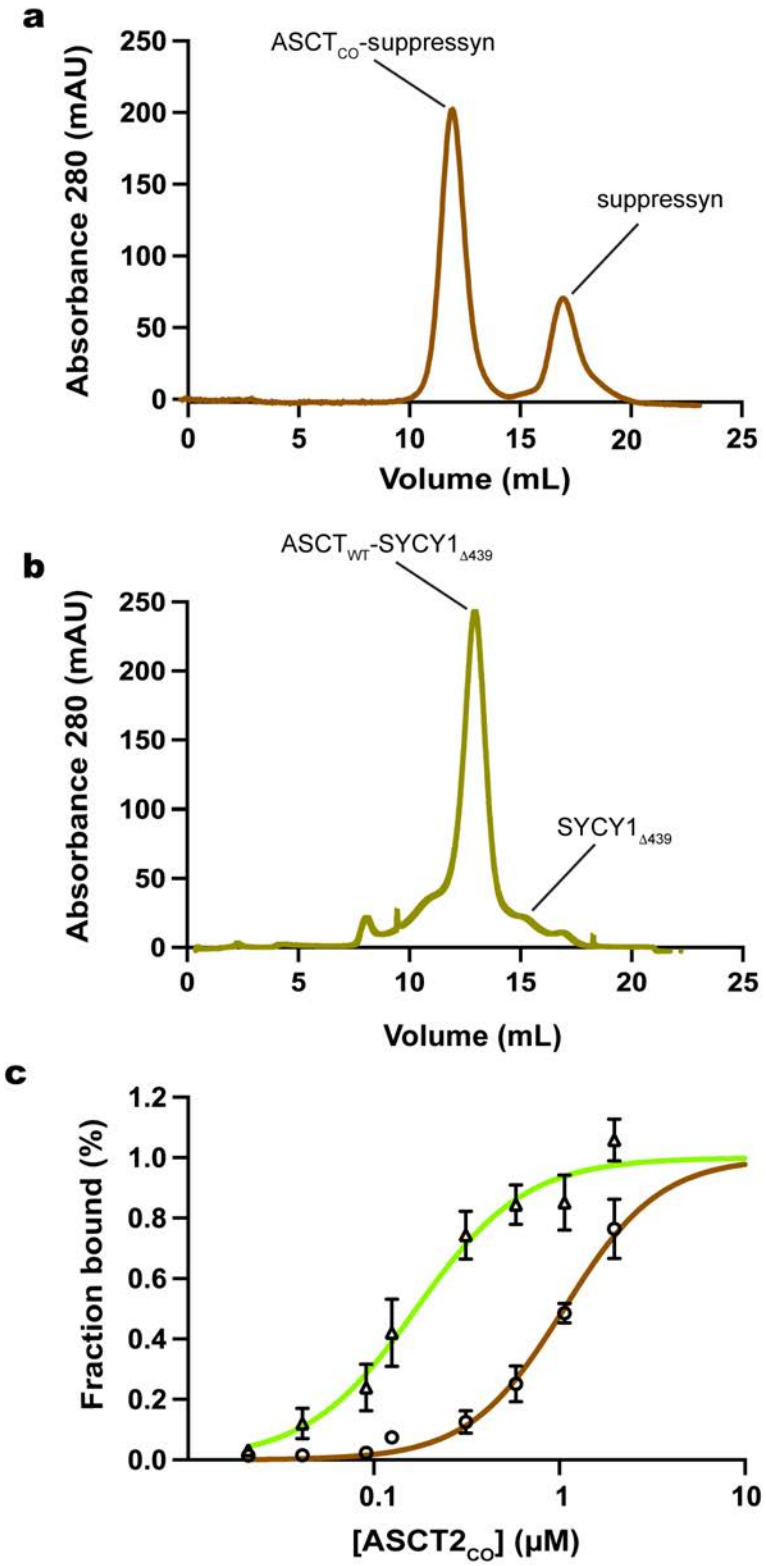
Figure 5: SYCY1$_{RBD}$ and SUPYN$_{RBD}$ inhibit ASCT2 transport function

a, Uptake of radioactive L-ala by ASCT2$_{CO}$ reconstituted in liposomes. **b**, Decrease in GFP-fluorescence after cleavage from the external side of the cytoplasmic GFP moiety fused to ASCT2$_{CO}$ in proteoliposomes. Inset depicts a cartoon representation of reconstituted transporters. **c**, Cytoplasmic view of the symmetric ASCT2$_{WT}$-Nb469 complex with the transporter represented as protein surface (tranDs in orange, and scaDs in blue) and the nanobody as ribbon (light green). **d**, Membrane view of an ASCT2$_{WT}$ subunit with the Nb469 bound at the tranD-scaD interface. The tranD is in an inward-facing state with the HP1 (red) fully exposed to the cytoplasmic bulk. **e**, Effects of SYCY1$_{Δ439}$ (left) and SUPYN$_{RBD}$ (right) on the Na^+-induced capacitive transients by ASCT2$_{CO}$ reconstituted in liposomes, in the presence of L-ala. The control experiments were done in the absence of HERV-derived proteins. Representative averaged current traces from three replicas are shown in the insets (middle). In a, b, and e bars depict averages of at least three biologically independent experiments, and error bars represent SEM. Black dots represent values from individual experiments.

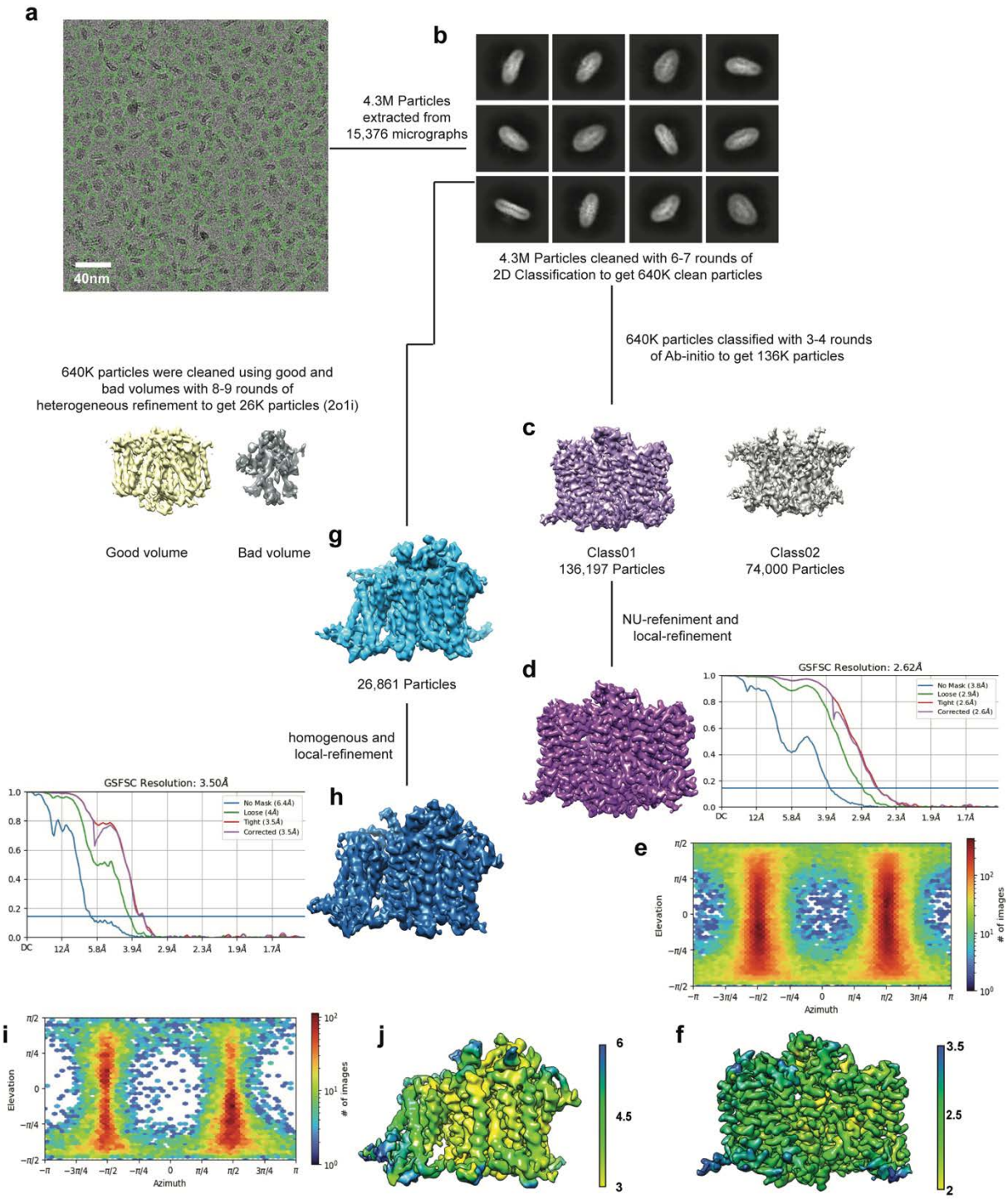


Extended Data Figure 1: ASCT2 consensus design

a, Amino acid sequence alignment of ASCT2$_{WT}$ and ASCT2$_{CO}$. **b**, Cell-cell fusion assay based on split-GFP. ASCT2$_{CO}$ shows similar levels of syncytin-1 mediated cell-cell fusion than ASCT2$_{WT}$. Purified SYCY1$_{\Delta 439}$ and suppressyn constructs recognize ASCT2$_{WT}$ and inhibit syncytin-1 mediated cell-cell fusion. The syncytin-1 mutant C186S precludes cell fusion and it is used as a negative control to account for background fluorescence. Bars depict averages of at least three biologically independent experiments, and error bars represent SEM. Black dots represent values from individual experiments. **c**, Reconstituted ASCT2$_{CO}$ yields nearly 5-fold higher radioactive substrate uptake in proteoliposomes compared to ASCT2$_{WT}$.

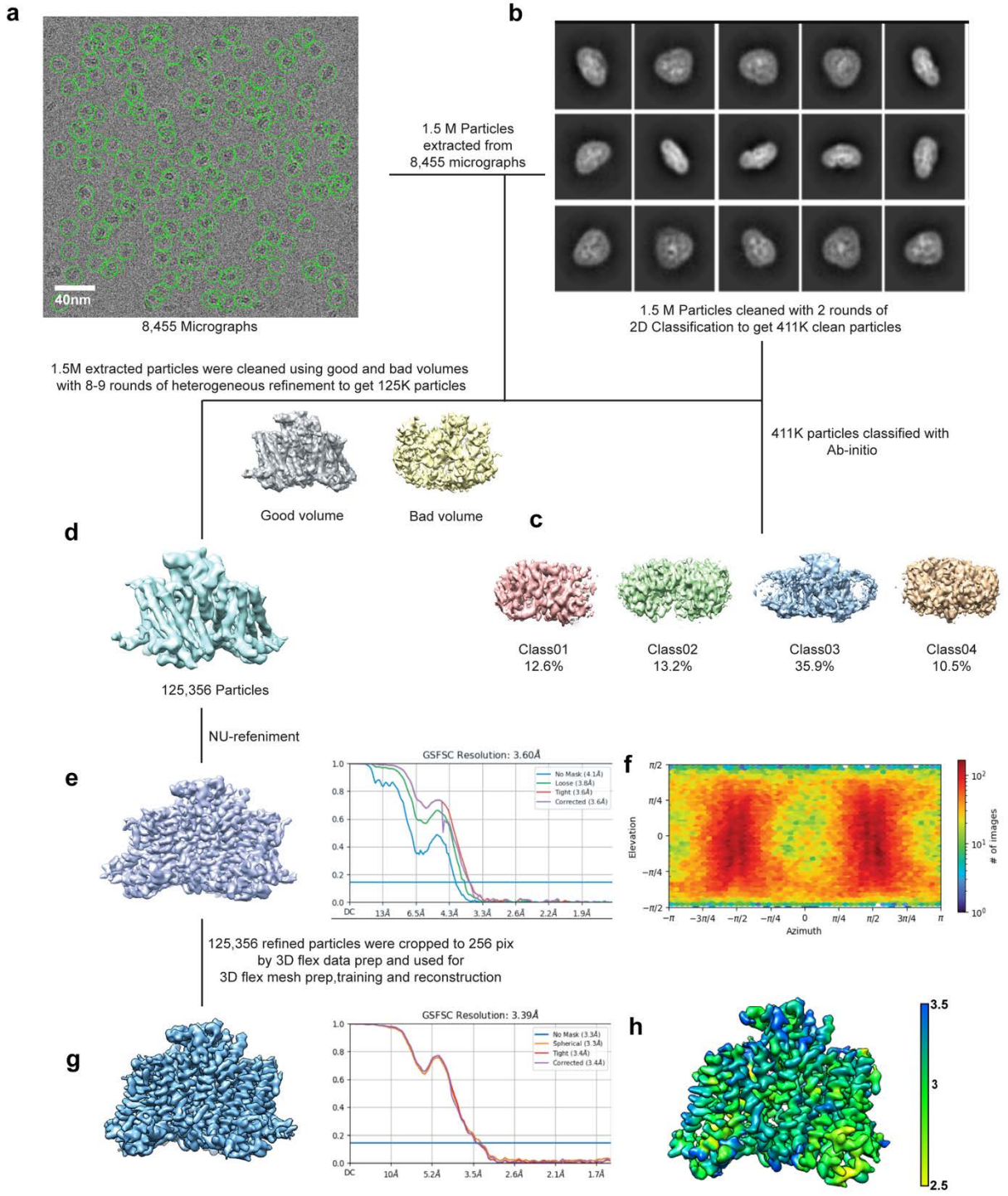


Extended Data Figure 2: ASCT2$_{WT}$-SYCY1-ΔG469 and ASCT2$_{CO}$-suppressyn complexes **a**, SEC profile of the ASCT2$_{CO}$-suppressyn complex. **b**, SEC profile of the ASCT2$_{WT}$-SYCY1$_{\Delta469}$ complex. **c**, Binding curves of suppressyn (triangles) and SYCY1$_{\Delta469}$ (circles) to ASCT2$_{CO}$ in detergent solution. Empty symbols depict the averages of 3 biologically independent experiments, and error bars represent SEM. Solid lines are fits of the Hill equation with $^{APP}K_D$ and Hill coefficient values of 1 μM and 1.7, respectively, for SYCY1$_{\Delta469}$, and 0.2 μM and 1.5, respectively, for suppressyn.

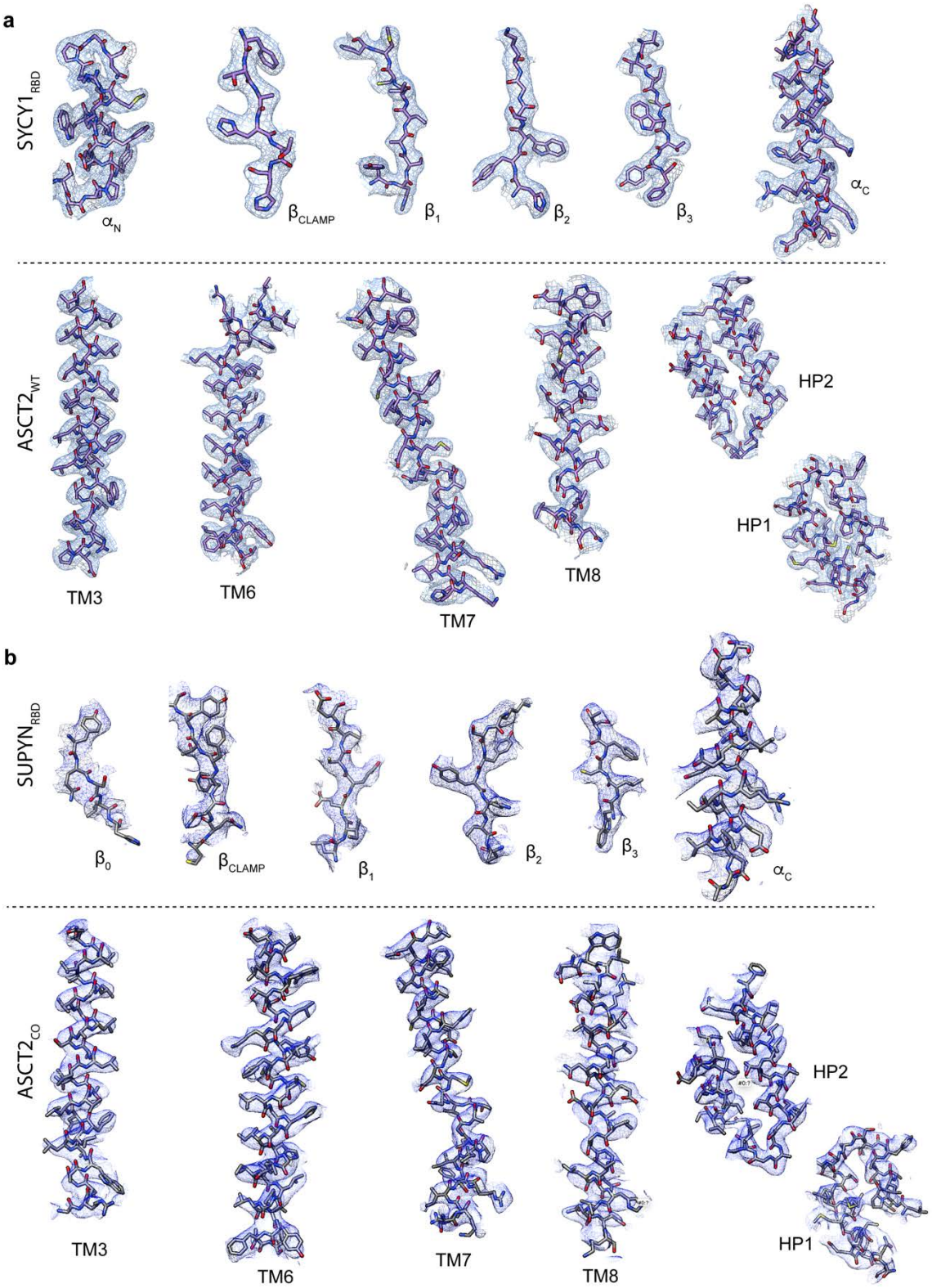


717

718 **Extended Data Figure 3: ASCT2$_{WT}$-SYCY1$_{RBD}$ cryo-EM data processing pipeline**
719 **a**, Representative EM micrograph. **b**, Gallery of representative 2D class-averages. **c**, 3D classes
720 from *ab initio* classification. **d**, Local refined map and Fourier Shell Correlation (FSC) plot
721 with FSC threshold at 0.143. **e**, Viewing direction distribution plot. **f**, The 2.6Å map (overall
722 resolution) is coloured based on local resolution with the scale bar in angstroms. **g**, 3D *ab-initio*
723 class after reprocessing 640k particles to improve the density of the third tranD. **h**, Local refined
724 map and corresponding FSC plot with threshold at 0.143. **i**, Viewing direction distribution plot.
725 **j**, The 3.5Å map (overall resolution) is coloured based on local resolution with the scale bar in
726 angstroms.
727

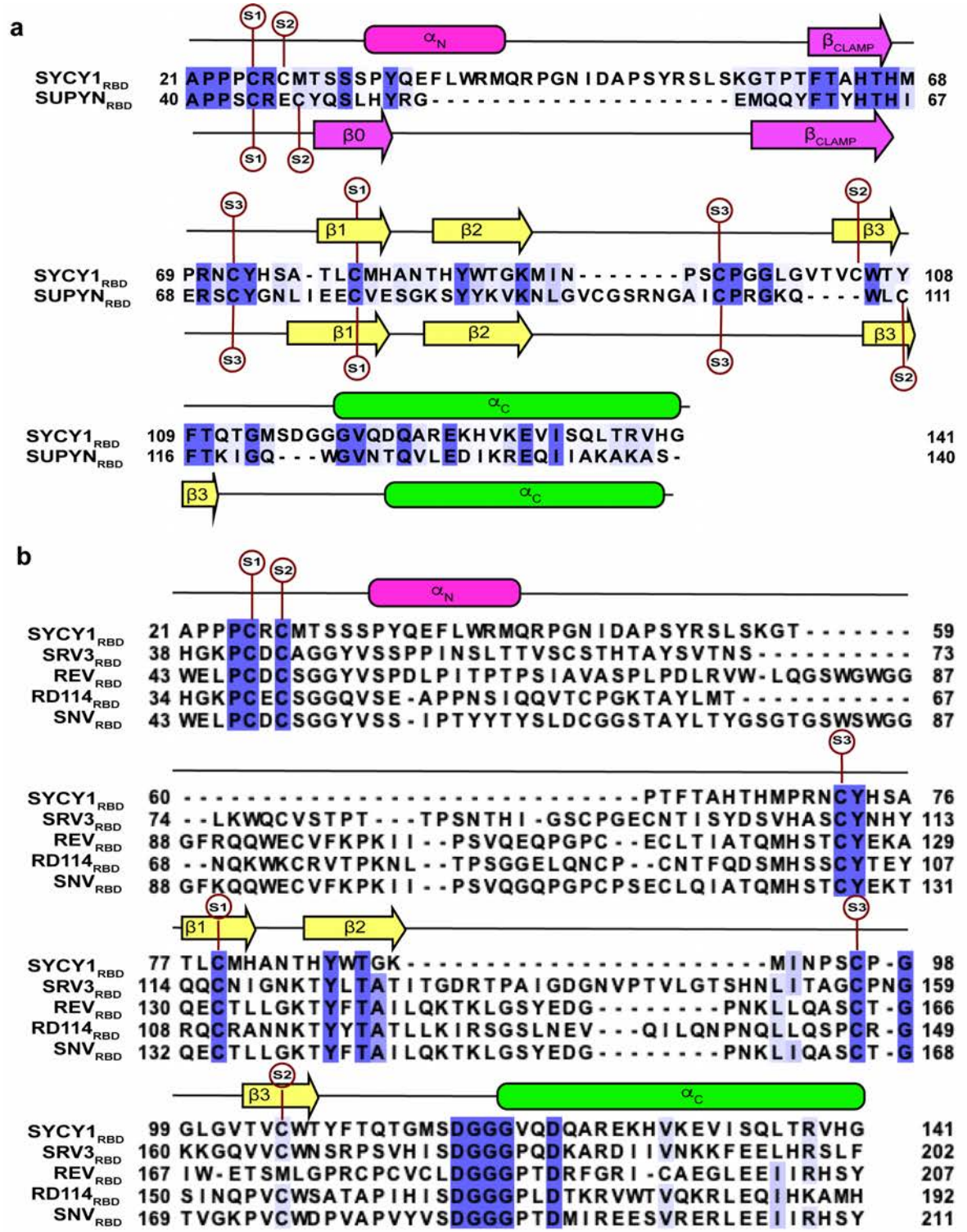


Extended Data Figure 4: ASCT2$_{CO}$-SUPYN$_{RBD}$ cryo-EM data processing pipeline
a, Representative EM micrograph. **b**, Gallery of representative 2D class-averages. **c**, 3D classes from *ab initio* classification. **d**, 3D class from heterogeneous refinement. **e**, Local refined map and Fourier Shell Correlation (FSC) plot with FSC threshold at 0.143. **f**, Viewing direction distribution plot. **g**, 3D-Flex reconstruction map and corresponding FSC plot with threshold at 0.143. **h**, The 3.6Å map (overall resolution) is coloured based on local resolution with the scale bar in angstroms.



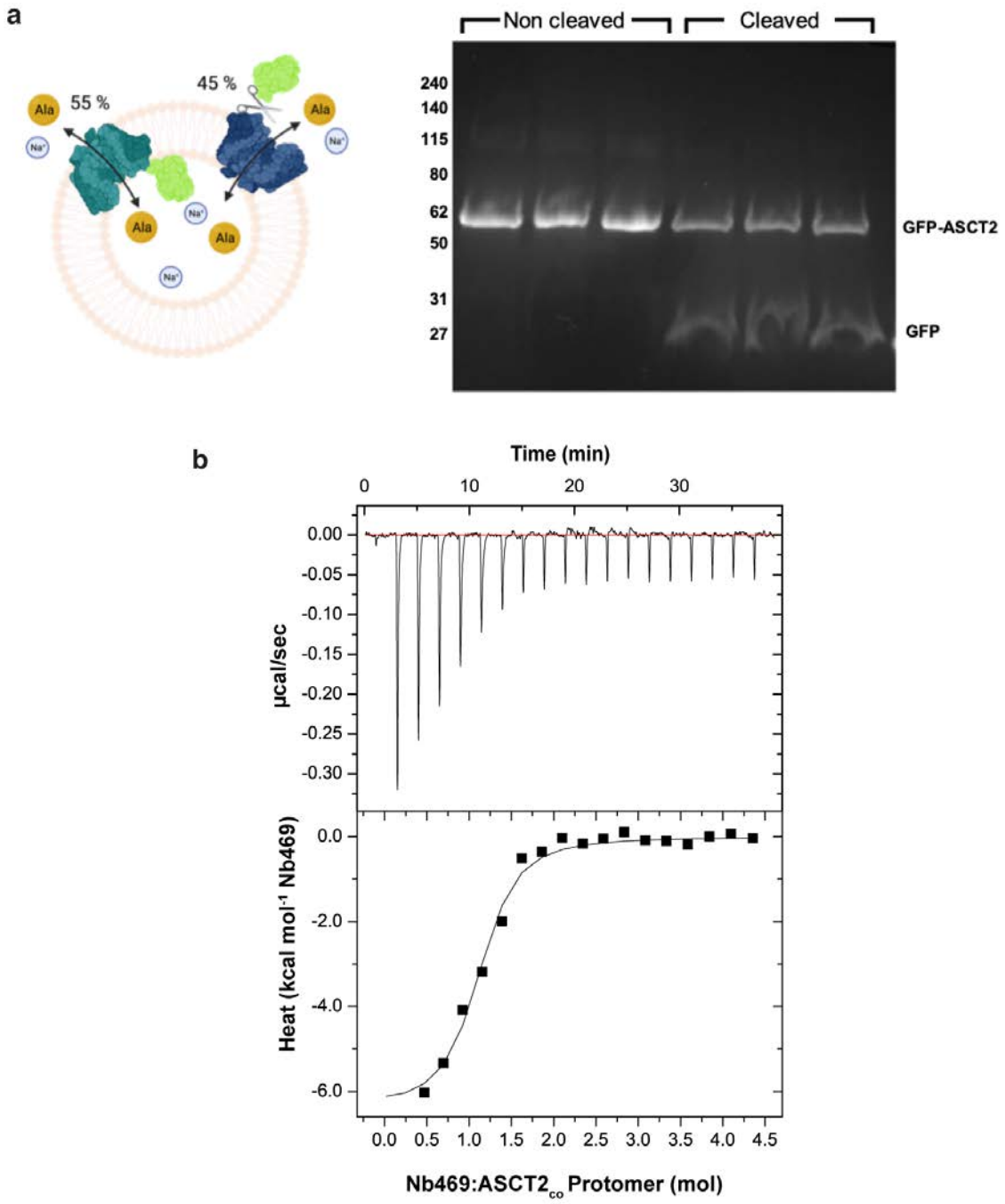
Extended Data Figure 5: Cryo-EM density of ASCT2$_{WT}$-SYCY1$_{RBD}$ and ASCT2$_{CO}$-SUPYN$_{RBD}$ complexes

a, Cryo-EM density corresponding to individual ASCT2$_{WT}$ and SYCY1$_{RBD}$ structural elements. **b**, Cryo-EM density corresponding to individual ASCT2$_{CO}$ and SUPYN$_{RBD}$ structural elements.

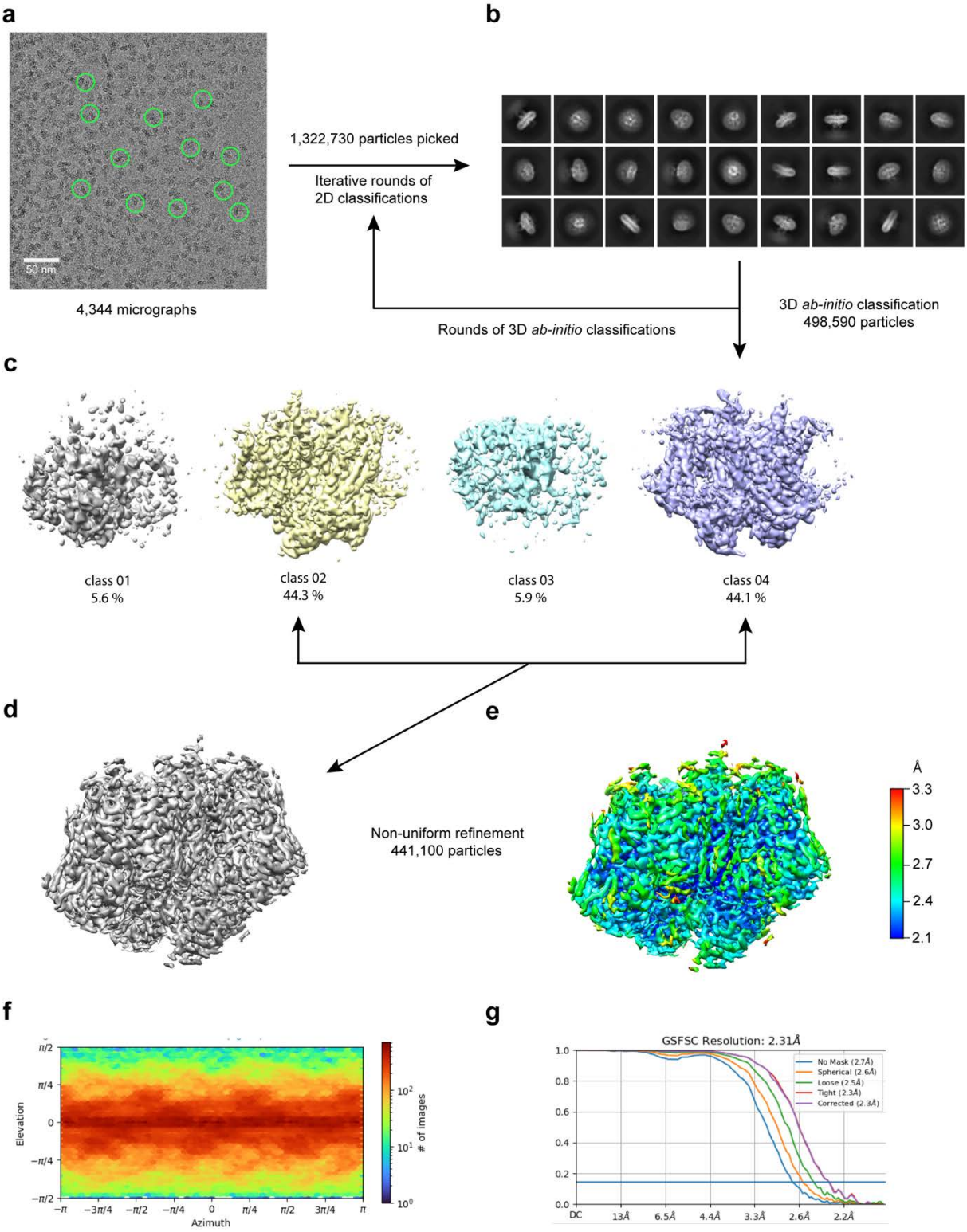


Extended Data Figure 6: Alignments of endogenous exogenous retroviral RBDs

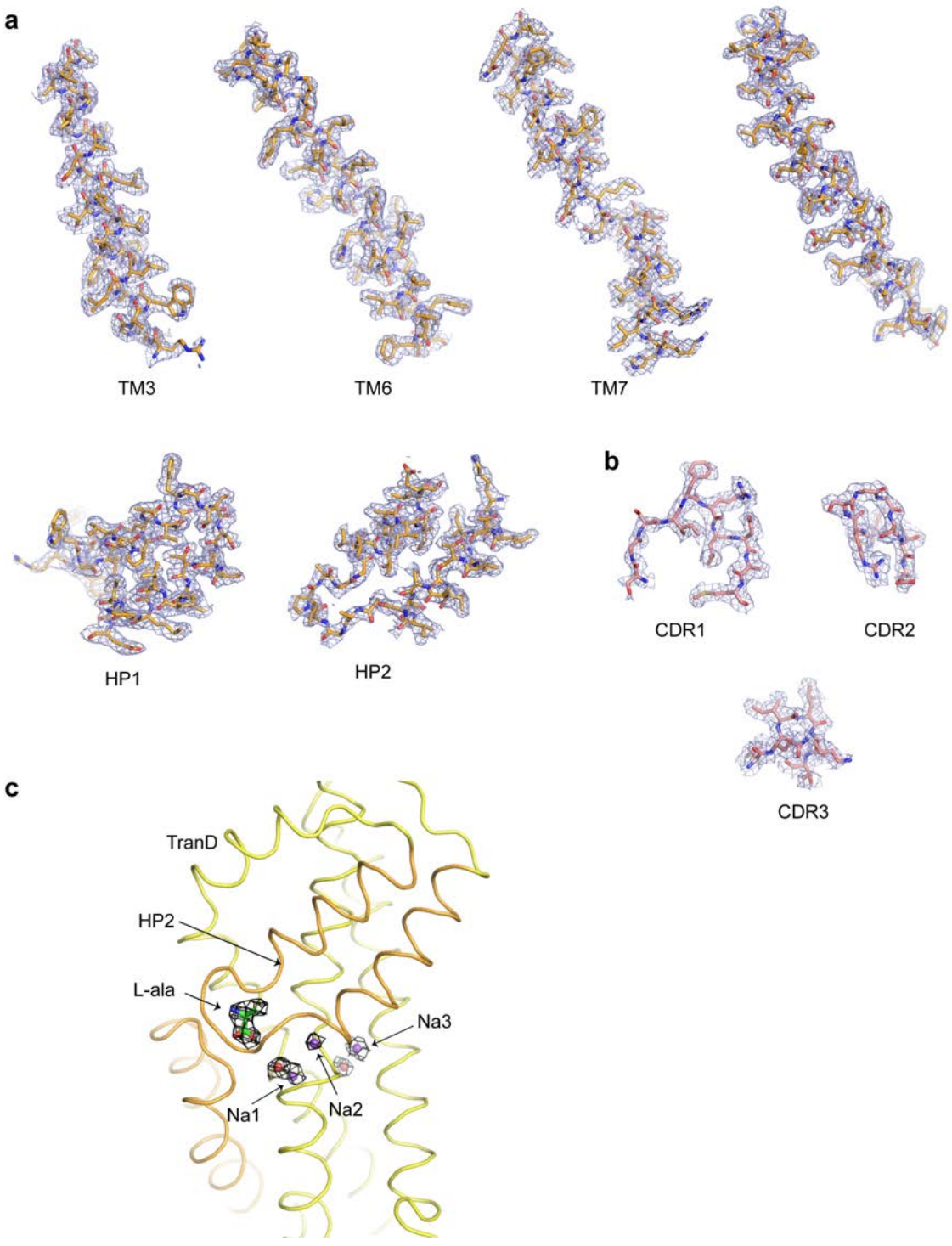
a, Sequence alignment of the $SYCY1_{RBD}$ and $SUPYN_{RBD}$. Disulfide bonds (S1, S2, and S3) and the secondary structural elements observed in the cryo-EM structures are marked above ($SYCY1_{RBD}$) and below ($SUPYN_{RBD}$) the alignment rows. **b**, Sequence alignment of the RBDs of simian retrovirus 3 ($SRV3_{RBD}$), avian reticuloendotheliosis virus (REV_{RBD}), simian type D retrovirus ($RD114_{RBD}$), avian spleen necrosis virus (SNV_{RBD}), and syncytin-1. Disulfide bonds (S1, S2, and S3) and the secondary structural elements observed in the $ASCT2_{WT}$-$SYCY1_{RBD}$ cryo-EM structure are marked.



752
753 **Extended Data Figure 7: Protease cleavage of GFP-ASCT2$_{CO}$ and Nb469 ITC binding**
754 **a**, Determination of ASCT2$_{CO}$ orientation in proteoliposomes. ASCT2$_{CO}$ N-terminal fused to
755 GFP was reconstituted in proteoliposomes and incubated in the presence or absence of 3C-
756 protease. In-gel fluorescence shows profiles of technical triplicates of non-cleaved and cleaved
757 (inside-out orientation) GFP-ASCT2$_{CO}$ in proteoliposomes. **b**, Isothermal titration calorimetry
758 analysis of Nb469 binding to ASTC$_{CO}$ at 25 ºC in detergent solutions. The black line in the
759 lower panel is the fit of the quadratic binding equation ("one site binding model") with the
760 following parameters: Nb469:ASCT$_{CO}$-protomer=1.1; K_D=360 nM; ΔH=6.3 kcal/mol; ΔS=8.0
761 cal/mol/deg.
762



Extended Data Figure 8: ASCT2$_{WT}$-Nb469 cryo-EM data processing pipeline
a, Representative EM micrograph with examples of individual picked particles (green circles). **b**, Gallery of representative 2D class-averages. **c**, 3D classes from *ab initio* classification. **d**, Non-uniform refinement map. **e**, The Non-uniform refinement map is coloured based on local resolution. **f**, Viewing direction distribution plot. **g**, Fourier Shell Correlation (FSC) plot of the non-uniform refinement with FSC threshold at 0.143.



Extended Data Figure 9: Cryo-EM density of ASCT2$_{WT}$-Nb469 complex
a, Cryo-EM density around different structural elements of ASCT2$_{WT}$. **b** Cryo-EM density around the complementary-determining regions of the Nb469. **c**, Cryo-EM density around the bound sodium ions (purple spheres), structural waters (red spheres), and L-ala substrate (green sticks). The tranD is shown as ribbon (colored in yellow) with the HP1 and HP2 colored in orange.

Extended Data Table 1: Cryo-EM data collection and processing

	$ASCT2_{WT}$-$SYCY1_{\Delta 439}$ (2.62Å)	$ASCT2_{WT}$- $SYCY1_{\Delta 439}$ (3.5Å)	$ASCT2_{CO}$-SUPYN (3.39Å)	ASCT2-Nb469 (2.31Å)
Data collection and processing				
Magnification	165,000	165,000	105,000	130,000
Voltage (kV)	300	300	300	300
Electron exposure ($e^{-}/Å^2$)	53.4	53.4	42.61	51.64
Defocus range (μm)	-0.4 to -1.5	-0.4 to -1.5	-0.8 to -2	-0.6 to -1.5
Pixel size (Å)	0.731	0.731	0.814	0.934
Symmetry imposed	None	None	None	C3
Initial particle images (no.)	4,018,074	500,000	1,540,255	1,322,730
Final particle images (no.)	136,197	26,861	125,356	441,100
Map resolution (Å)	2.62	3.5	3.39	2.31
FSC threshold	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.6-3.5	3.5-7	3-6	2.3-3.5
Refinement				
Initial model used (PDB code)	6GCT	6GCT and 6MPB	6GCT	6GCT and 7PQQ
Model resolution	2.62	3.5	3.39	2.31
FSC threshold	0.143	0.143	0.143	0.143
Model resolution range	2.6-3.5	3.5-7	3-6	2.3-3.5
Map sharpening B-factor ($Å^2$)	-60.1	-82.2	-90.0	-77.4
Model composition				
Non-hydrogen atoms	8,922	10,620	10,381	12,126
Protein residues	1,190	1,420	1,381	1,596
Ligand	2	2	2	12
Waters	-	-	-	93
B factors ($Å^2$)				
Protein	104.24	79.70	108.83	33.85
Ligand	117.07	74.98	116.07	25.82
Waters	-	-	-	28.80
R.m.s. deviations				
Bond lengths (Å)	0.003	0.006	0.004	0.004
Bond angles (°)	0.628	1.225	0.728	0.588
Validation				
MolProbity score	1.61	2.22	1.79	1.20
Clashscore	5.88	9.77	9.52	4.23
Poor rotamers (%)	0.00	2.93	0.27	0.48
Ramachandran plot				
Favored (%)	95.83	94.88	95.81	99.62
Allowed (%)	4.17	4.84	4.19	0.38

References

1 Mi, S. *et al.* Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* **403**, 785-789, doi:10.1038/35001608 (2000).

2 Blond, J. L. *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* **74**, 3321-3329, doi:10.1128/jvi.74.7.3321-3329.2000 (2000).

3 Mallet, F. *et al.* The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology. *Proc Natl Acad Sci U S A* **101**, 1731-1736, doi:10.1073/pnas.0305763101 (2004).

4 Sugimoto, J., Sugimoto, M., Bernstein, H., Jinno, Y. & Schust, D. A novel human endogenous retroviral protein inhibits cell-cell fusion. *Sci Rep* **3**, 1462, doi:10.1038/srep01462 (2013).

5 Sugimoto, J. *et al.* Suppressyn localization and dynamic expression patterns in primary human tissues support a physiologic role in human placentation. *Sci Rep* **9**, 19502, doi:10.1038/s41598-019-55933-x (2019).

6 Rasko, J. E., Battini, J. L., Gottschalk, R. J., Mazo, I. & Miller, A. D. The RD114/simian type D retrovirus receptor is a neutral amino acid transporter. *Proc Natl Acad Sci U S A* **96**, 2129-2134, doi:10.1073/pnas.96.5.2129 (1999).

7 Tailor, C. S., Nouri, A., Zhao, Y., Takeuchi, Y. & Kabat, D. A sodium-dependent neutral-amino-acid transporter mediates infections of feline and baboon endogenous retroviruses and simian type D retroviruses. *J Virol* **73**, 4470-4474, doi:10.1128/JVI.73.5.4470-4474.1999 (1999).

8 Sommerfelt, M. A. & Weiss, R. A. Receptor interference groups of 20 retroviruses plating on human cells. *Virology* **176**, 58-69, doi:10.1016/0042-6822(90)90230-o (1990).

9 Ponferrada, V. G., Mauck, B. S. & Wooley, D. P. The envelope glycoprotein of human endogenous retrovirus HERV-W induces cellular resistance to spleen necrosis virus. *Arch Virol* **148**, 659-675, doi:10.1007/s00705-002-0960-x (2003).

10 Frank, J. A. *et al.* Evolution and antiviral activity of a human protein of retroviral origin. *Science* **378**, 422-428, doi:10.1126/science.abq7871 (2022).

11 Robbins, J. R. & Bakardjiev, A. I. Pathogens and the placental fortress. *Curr Opin Microbiol* **15**, 36-43, doi:10.1016/j.mib.2011.11.006 (2012).

12 Lavialle, C. *et al.* Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation. *Philos Trans R Soc Lond B Biol Sci* **368**, 20120507, doi:10.1098/rstb.2012.0507 (2013).

13 Blond, J. L. *et al.* Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family. *J Virol* **73**, 1175-1185, doi:10.1128/JVI.73.2.1175-1185.1999 (1999).

14 Langbein, M. *et al.* Impaired cytotrophoblast cell-cell fusion is associated with reduced Syncytin and increased apoptosis in patients with placental dysfunction. *Mol Reprod Dev* **75**, 175-183, doi:10.1002/mrd.20729 (2008).

15 Lee, X. *et al.* Downregulation of placental syncytin expression and abnormal protein localization in pre-eclampsia. *Placenta* **22**, 808-812, doi:10.1053/plac.2001.0722 (2001).

16 Bolze, P. A., Mommert, M. & Mallet, F. Contribution of Syncytins and Other Endogenous Retroviral Envelopes to Human Placenta Pathologies. *Prog Mol Biol Transl Sci* **145**, 111-162, doi:10.1016/bs.pmbts.2016.12.005 (2017).
17 Cheynet, V. *et al.* Synthesis, assembly, and processing of the Env ERVWE1/syncytin human endogenous retroviral envelope. *J Virol* **79**, 5585-5593, doi:10.1128/JVI.79.9.5585-5593.2005 (2005).
18 Chang, C., Chen, P. T., Chang, G. D., Huang, C. J. & Chen, H. Functional characterization of the placental fusogenic membrane protein syncytin. *Biol Reprod* **71**, 1956-1962, doi:10.1095/biolreprod.104.033340 (2004).
19 Lavillette, D. *et al.* The envelope glycoprotein of human endogenous retrovirus type W uses a divergent family of amino acid transporters/cell surface receptors. *J Virol* **76**, 6442-6452, doi:10.1128/jvi.76.13.6442-6452.2002 (2002).
20 Hogan, V. & Johnson, W. E. Unique Structure and Distinctive Properties of the Ancient and Ubiquitous Gamma-Type Envelope Glycoprotein. *Viruses* **15**, doi:10.3390/v15020274 (2023).
21 Henzy, J. E. & Johnson, W. E. Pushing the endogenous envelope. *Philos Trans R Soc Lond B Biol Sci* **368**, 20120506, doi:10.1098/rstb.2012.0506 (2013).
22 Ruigrok, K. *et al.* X-ray Structures of the Post-fusion 6-Helix Bundle of the Human Syncytins and their Functional Implications. *J Mol Biol* **431**, 4922-4940, doi:10.1016/j.jmb.2019.10.020 (2019).
23 Sugimoto, J. *et al.* Could the Human Endogenous Retrovirus-Derived Syncytialization Inhibitor, Suppressyn, Limit Heterotypic Cell Fusion Events in the Decidua? *Int J Mol Sci* **22**, doi:10.3390/ijms221910259 (2021).
24 Sinha, A. & Johnson, W. E. Retroviruses of the RDR superinfection interference group: ancient origins and broad host distribution of a promiscuous Env gene. *Curr Opin Virol* **25**, 105-112, doi:10.1016/j.coviro.2017.07.020 (2017).
25 Stromberg, K. *et al.* Characterization of exogenous type D retrovirus from a fibroma of a macaque with simian AIDS and fibromatosis. *Science* **224**, 289-282, doi:10.1126/science.6200929 (1984).
26 Daniel, M. D. *et al.* A new type D retrovirus isolated from macaques with an immunodeficiency syndrome. *Science* **223**, 602-605, doi:10.1126/science.6695172 (1984).
27 Grange, Z. L. *et al.* Ranking the risk of animal-to-human spillover for newly discovered viruses. *Proc Natl Acad Sci U S A* **118**, doi:10.1073/pnas.2002324118 (2021).
28 Wozniakowski, G., Frant, M. & Mamczur, A. Avian Reticuloendotheliosis in Chickens - An Update on Disease Occurrence and Clinical Course. *J Vet Res* **62**, 257-260, doi:10.2478/jvetres-2018-0036 (2018).
29 Cheynet, V., Oriol, G. & Mallet, F. Identification of the hASCT2-binding domain of the Env ERVWE1/syncytin-1 fusogenic glycoprotein. *Retrovirology* **3**, 41, doi:10.1186/1742-4690-3-41 (2006).
30 Kanai, Y. *et al.* The SLC1 high-affinity glutamate and neutral amino acid transporter family. *Mol Aspects Med* **34**, 108-120, doi:10.1016/j.mam.2013.01.001 (2013).
31 Wahi, K. & Holst, J. ASCT2: a potential cancer drug target. *Expert Opin Ther Targets* **23**, 555-558, doi:10.1080/14728222.2019.1627328 (2019).
32 Liu, Y. *et al.* The role of ASCT2 in cancer: A review. *Eur J Pharmacol* **837**, 81-87, doi:10.1016/j.ejphar.2018.07.007 (2018).

33 Garaeva, A. A. *et al.* Cryo-EM structure of the human neutral amino acid transporter ASCT2. *Nat Struct Mol Biol* **25**, 515-521, doi:10.1038/s41594-018-0076-y (2018).

34 Yu, X. *et al.* Cryo-EM structures of the human glutamine transporter SLC1A5 (ASCT2) in the outward-facing conformation. *Elife* **8**, doi:10.7554/eLife.48120 (2019).

35 Yernool, D., Boudker, O., Jin, Y. & Gouaux, E. Structure of a glutamate transporter homologue from Pyrococcus horikoshii. *Nature* **431**, 811-818, doi:10.1038/nature03018 (2004).

36 Canul-Tec, J. C. *et al.* Structure and allosteric inhibition of excitatory amino acid transporter 1. *Nature* **544**, 446-451, doi:10.1038/nature22064 (2017).

37 Reyes, N., Ginter, C. & Boudker, O. Transport mechanism of a bacterial homologue of glutamate transporters. *Nature* **462**, 880-885, doi:10.1038/nature08616 (2009).

38 Crisman, T. J., Qu, S., Kanner, B. I. & Forrest, L. R. Inward-facing conformation of glutamate transporters as revealed by their inverted-topology structural repeats. *Proc Natl Acad Sci U S A* **106**, 20752-20757, doi:10.1073/pnas.0908570106 (2009).

39 Marin, M., Lavillette, D., Kelly, S. M. & Kabat, D. N-linked glycosylation and sequence changes in a critical negative control region of the ASCT1 and ASCT2 neutral amino acid transporters determine their retroviral receptor functions. *J Virol* **77**, 2936-2945, doi:10.1128/jvi.77.5.2936-2945.2003 (2003).

40 Cirri, E. *et al.* Consensus designs and thermal stability determinants of a human glutamate transporter. *Elife* **7**, doi:10.7554/eLife.40110 (2018).

41 Martinez-Molledo, M., Nji, E. & Reyes, N. Structural insights into the lysophospholipid brain uptake mechanism and its inhibition by syncytin-2. *Nat Struct Mol Biol* **29**, 604-612, doi:10.1038/s41594-022-00786-8 (2022).

42 Fass, D. *et al.* Structure of a murine leukemia virus receptor-binding glycoprotein at 2.0 angstrom resolution. *Science* **277**, 1662-1666, doi:10.1126/science.277.5332.1662 (1997).

43 Barnett, A. L., Wensel, D. L., Li, W., Fass, D. & Cunningham, J. M. Structure and mechanism of a coreceptor for infection by a pathogenic feline retrovirus. *J Virol* **77**, 2717-2729, doi:10.1128/jvi.77.4.2717-2729.2003 (2003).

44 McCarthy, K. R. *et al.* Structure of the Receptor Binding Domain of EnvP(b)1, an Endogenous Retroviral Envelope Protein Expressed in Human Tissues. *mBio* **11**, doi:10.1128/mBio.02772-20 (2020).

45 Holm, L. DALI and the persistence of protein shape. *Protein Sci* **29**, 128-140, doi:10.1002/pro.3749 (2020).

46 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).

47 Hotzel, I. Deep-Time Structural Evolution of Retroviral and Filoviral Surface Envelope Proteins. *J Virol* **96**, e0006322, doi:10.1128/jvi.00063-22 (2022).

48 Boudker, O., Ryan, R. M., Yernool, D., Shimamoto, K. & Gouaux, E. Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. *Nature* **445**, 387-393, doi:10.1038/nature05455 (2007).

49 Garaeva, A. A., Guskov, A., Slotboom, D. J. & Paulino, C. A one-gate elevator mechanism for the human neutral amino acid transporter ASCT2. *Nat Commun* **10**, 3427, doi:10.1038/s41467-019-11363-x (2019).

50 Benton, D. J. *et al.* Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* **588**, 327-330, doi:10.1038/s41586-020-2772-0 (2020).

51 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* **100**, 13013-13018, doi:10.1073/pnas.2132646100 (2003).

52 Esnault, C. *et al.* A placenta-specific receptor for the fusogenic, endogenous retrovirus-derived, human syncytin-2. *Proc Natl Acad Sci U S A* **105**, 17532-17537, doi:10.1073/pnas.0807413105 (2008).

53 Shaik, M. M. *et al.* Structural basis of coreceptor recognition by HIV-1 envelope spike. *Nature* **565**, 318-323, doi:10.1038/s41586-018-0804-9 (2019).

54 Gong, X. *et al.* Structural Insights into the Niemann-Pick C1 (NPC1)-Mediated Cholesterol Transfer and Ebola Infection. *Cell* **165**, 1467-1478, doi:10.1016/j.cell.2016.05.022 (2016).

55 Yan, R. *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444-1448, doi:10.1126/science.abb2762 (2020).

56 Garibsingh, R. A. *et al.* Rational design of ASCT2 inhibitors using an integrated experimental-computational approach. *Proc Natl Acad Sci U S A* **118**, doi:10.1073/pnas.2104093118 (2021).

57 Iwaki, T., Figuera, M., Ploplis, V. A. & Castellino, F. J. Rapid selection of Drosophila S2 cells with the puromycin resistance gene. *Biotechniques* **35**, 482-484, 486, doi:10.2144/03353bm08 (2003).

58 Backovic, M. & Krey, T. Stable Drosophila Cell Lines: An Alternative Approach to Exogenous Protein Expression. *Methods Mol Biol* **1350**, 349-358, doi:10.1007/978-1-4939-3043-2_17 (2016).

59 Pardon, E. *et al.* A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc* **9**, 674-693, doi:10.1038/nprot.2014.039 (2014).

60 Cabantous, S., Terwilliger, T. C. & Waldo, G. S. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat Biotechnol* **23**, 102-107, doi:10.1038/nbt1044 (2005).

61 Guo, H. *et al.* Electron-event representation data enable efficient cryoEM file storage with full preservation of spatial and temporal resolution. *IUCrJ* **7**, 860-869, doi:10.1107/S205225252000929X (2020).

62 Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51, doi:10.1016/j.jsb.2005.07.007 (2005).

63 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).

64 Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat Methods* **17**, 1214-1221, doi:10.1038/s41592-020-00990-8 (2020).

65 Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun Biol* **4**, 874, doi:10.1038/s42003-021-02399-1 (2021).

66 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

67 Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444, doi:10.1093/nar/gkab1061 (2022).
68 Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153-1160, doi:10.1038/s41592-019-0575-8 (2019).
69 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).
70 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).
71 Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* **19**, 679-682, doi:10.1038/s41592-022-01488-1 (2022).
72 Pavic, A., Holmes, A. O. M., Postis, V. L. G. & Goldman, A. Glutamate transporters: a broad review of the most recent archaeal and human structures. *Biochem Soc Trans* **47**, 1197-1207, doi:10.1042/BST20190316 (2019).
73 Danbolt, N. C. Glutamate uptake. *Prog Neurobiol* **65**, 1-105, doi:10.1016/s0301-0082(00)00067-8 (2001).
74 Guskov, A., Jensen, S., Faustino, I., Marrink, S. J. & Slotboom, D. J. Coupled binding mechanism of three sodium ions and aspartate in the glutamate transporter homologue GltTk. *Nat Commun* **7**, 13420, doi:10.1038/ncomms13420 (2016).
75 Canul-Tec, J. C. *et al.* The ion-coupling mechanism of human excitatory amino acid transporters. *EMBO J* **41**, e108341, doi:10.15252/embj.2021108341 (2022).