Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses

Jonathan Abraham[1,2], Kevin D Corbett[3], Michael Farzan[4], Hyeryun Choe[2] & Stephen C Harrison[1,3,5]

New World hemorrhagic fever arenaviruses are rodent-borne agents that cause severe human disease. The GP1 subunit of the surface glycoprotein mediates cell attachment through transferrin receptor 1 (TfR1). We report the structure of Machupo virus (MACV) GP1 bound with human TfR1. Atomic details of the GP1-TfR1 interface clarify the importance of TfR1 residues implicated in New World arenavirus host specificity. Analysis of sequence variation among New World arenavirus GP1s and their host-species receptors, in light of the molecular structure, indicates determinants of viral zoonotic transmission. Infectivities of pseudoviruses in cells expressing mutated TfR1 confirm that contacts at the tip of the TfR1 apical domain determine the capacity of human TfR1 to mediate infection by particular New World arenaviruses. We propose that New World arenaviruses that are pathogenic to humans fortuitously acquired affinity for human TfR1 during adaptation to TfR1 of their natural hosts.

Arenaviruses have enveloped virions and single-stranded, bisegmented, ambisense RNA genomes[1]. They fall into two groups, the 'Old World' and 'New World' arenaviruses; the latter are further divided into clades A, B, C and a recombinant A/B clade[2–4]. Several New World clade B arenaviruses cause hemorrhagic fevers when transmitted from their host species to humans. These include the Machupo (MACV), Junín (JUNV), Guanarito (GTOV) and Sabiá (SABV) viruses, the agents of Bolivian, Argentine, Venezuelan and Brazilian hemorrhagic fevers, respectively, all with high case fatality rates[5–7]. Chapare virus is a novel New World clade B hemorrhagic fever arenavirus recently isolated in Bolivia[8].

The arenavirus envelope glycoprotein GP is the sole protein on the virion surface. During its maturation, GP is processed into three noncovalently associated subunits: the stable signal peptide (SSP), GP1 and GP2 (ref. 9). Unlike most signal peptides, SSP remains on the surface of virions and participates in maturation and pH-dependent membrane-fusion activity of the GP complex[10–13]. The GP1 subunit interacts with cellular receptor(s), and GP2 mediates membrane fusion after virus particles are internalized into acidified endosomes[14–16].

The cellular receptor for the New World hemorrhagic fever arenaviruses MACV, JUNV, GTOV and SABV is transferrin receptor 1 (TfR1)[17–19]. TfR1 is an ubiquitously expressed, homodimeric, type II transmembrane glycoprotein that assists iron uptake into cells through endocytosis of iron-loaded transferrin[20]. The dimeric ectodomain of human TfR1 (residues 89–760) is a butterfly-shaped molecule with three subdomains: a 'protease-like' domain related in structure to certain zinc metalloproteases, an 'apical domain' and a 'helical domain', which makes the principal dimer contacts[21].

Transferrin interacts with the protease-like and helical domains[22] but does not interfere with the entry of MACV[17]. Mutational analysis implicates the TfR1 apical domain as the principal site of interaction with New World arenaviral GPs, and a short segment in the TfR1 apical domain (residues 208–212) is a critical determinant of virus host specificity[18,23].

TfR1 orthologs from the natural host species of all clade B New World arenaviruses we have tested (MACV, JUNV, GTOV, AMAV and TCRV) are functional cellular receptors for the corresponding virus, but human TfR1 (hTfR1) mediates entry only of the human pathogens[17–19,23] (**Table 1**). Thus, there is a strong correlation between the capacity of hTfR1 to mediate infection and the capacity of a particular New World clade B arenavirus to cause human disease[19,23].

To examine the structural basis for hTfR1 recognition by New World clade B arenaviruses, we have determined the crystal structure of MACV GP1 in complex with the ectodomain of hTfR1. As anticipated, MACV GP1 binds TfR1 through an extensive network of contacts with the apical domain. The atomic details of the GP1-TfR1 interface clarify the importance of several receptor residues implicated in New World arenavirus host specificity. On the viral glycoprotein side of this interface, there is substantial residue-sequence variability among GP1 proteins that bind the human receptor. We correlate this variability with differences among the sequences of the host-species TfR1 orthologs, which can serve as efficient receptors for the corresponding virus[18,23]. The specificities of the viral glycoproteins reflect long-term adaptation of the viruses in their natural hosts[24]; our data suggest that the New World hemorrhagic fever arenaviruses independently and fortuitously acquired their ability to bind human TfR1 and hence to cause human disease.

---

[1]Laboratory of Molecular Medicine and [2]Department of Medicine, Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA. [3]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA. [4]Department of Microbiology and Molecular Genetics, Harvard Medical School, New England Primate Center, Southborough, Massachusetts, USA. [5]Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts, USA. Correspondence should be addressed to S.C.H. (harrison@crystal.harvard.edu).

Table 1 Receptor usage by New World clade B arenaviruses

	*Cc*TfR1	*Cm*TfR1	*Aj*TfR1	*Zb*TfR1	*Ns*TfR1	*h*TfR1
MACV	✓		✓	✓	✓	✓
JUNV	✓	✓	✓	✓	✓	✓
TCRV			✓	✓	✓	
GTOV				✓	✓	✓
AMAV				✓	✓	
SABV	–	–	–	–	–	✓

Derived from refs. 17–19,23. Check marks indicate that the virus can infect cells expressing TfR1 derived from the host species indicated in the top row. The host species for SABV is unknown. Dashes indicate virus-receptor pairings not tested. Abbreviations: *Cc*, Calomys callosus; *Cm*, Calomys musculinus; *Aj*, Artibeus jamaicensis; *Zb*, Zygodontomys brevicauda; *Ns*, Neacomys spinosus; h, human. Viruses and their host-species receptors are color matched, with the exception of SABV.

RESULTS

Structure determination

We expressed recombinant MACV GP1$_{79-258}$ in insect cells with a hexahistidine tag at the C terminus, purified it using nickel-affinity chromatography and then treated it with elastase to remove the hexahistidine tag and five N-terminal residues and further purified it by gel filtration. We used a stably expressing Chinese hamster ovary (CHO) cell line to express a soluble form of the ectodomain of hTfR1 (residues 117–760), which we purified as described previously[21]. By SDS-PAGE analysis, the MACV GP1 and hTfR1 proteins migrated with molecular masses consistent with their full glycosylation (**Supplementary Fig. 1**). The proteins, mixed in equimolar ratio, crystallized in the space group $C222_1$ (**Table 2**). We determined the structure of the GP1–hTfR1 complex using molecular replacement, with the hTfR1 monomer[21] (PDB 1CX8) serving as a search model, and refined it using data extending to 2.4 Å (but incomplete beyond 2.7 Å: **Table 2**). The final model of the GP1–hTfR1 complex contains residues 121–756 of the receptor, except for residues 316–319, which are disordered, and residues 86–241 of MACV GP1, with N-linked glycans at TfR1 positions 251 and 727 and at GP1 positions 95, 137, 166 and 178.

Complex of MACV GP1 with human TfR1

MACV GP1 contacts the tip of the apical domain of TfR1, with 1,900 Å² of total surface area buried at the interface (**Fig. 1**). The apical domain of TfR1 is an insertion between the first and second strands in the central β-sheet of the protease-like domain. It resembles a β-sandwich in which the two sheets are splayed apart, with a helix (αII-2) running along the open edge, a fold similar to domain 4 of aconitase[21]. Helix αII-2 and strand βII-2 in the apical domain form most of the contacts with MACV GP1 (**Fig. 1b**). Residues 208–212, built as an extended loop in previous structures of hTfR1 (refs. 21,25), have a β-strand conformation in our structure of the complex and form an extension of strand βII-2. This segment was poorly defined in previous structures of hTfR1 (refs. 21,25), which were determined at lower resolution (3.1 Å and 2.9 Å), and further analysis of the electron density in the unliganded hTfR1 structure[21] (PDB 1CX8) in light of our higher-resolution model shows that residues 208–212 in unliganded hTfR1 adopt a very similar β-strand conformation (**Supplementary Fig. 2a**). Thus, GP1 binding induces no major main chain distortions in the TfR1 apical domain. It does, however, increase order in the extended βII-2 strand, displace the

backbone, and change the orientation of Leu209 and switch the rotameric state of Tyr211 (**Supplementary Fig. 2b**).

The core of MACV GP1 is a seven-stranded antiparallel β-sheet (β1–β7) with five interconnecting α-helical segments (αA–αE)[26] (**Fig. 2**). Four N-linked glycans decorate surfaces of the viral glycoprotein not involved in receptor binding. We could interpret density for seven sugar residues on Asn178, three (including a fucosyl group) on Asn166 and a single *N*-acetylglucosamine at the remaining two positions (Asn95 and Asn137). Several extended loops protrude from the exposed, concave face of the MACV GP1 core to contact TfR1 (**Fig. 2c**). The loop between β-strands 3 and 4 (loop 3, residues 113–123) contacts the βII-2 strand in the apical domain of hTfR1; a second loop (loop 10, residues 220–234), which is disulfide-linked and contains a short α-helical segment, packs against the opposite side of the βII-2 strand. Loop 10 is longer in MACV (15 residues) than in the GPs of other New World arenaviruses, whose GP1s have only a short segment (5 or 6 residues, **Fig. 2b**). A third loop (loop 7, residues 165–174, between β-strands 5 and 6), containing the glycan at Asn166, also contacts TfR1 (**Fig. 2c**).

MACV GP1–hTfR1 interface

We have analyzed the interface between MACV GP1 and the apical domain of TfR1 in terms of five interaction motifs, each anchored by a key residue on the receptor surface. Three of these five TfR1 residues—Tyr211, Asn348 and Val210—have previously been implicated as determinants of New World arenavirus host specificity[18,23].

Table 2 Data collection and refinement statistics (molecular replacement)

	MACV GP1:hTfR1
Data collection	
Space group	$C222_1$
Cell dimensions	
a, *b*, *c* (Å)	144.0, 170.2, 98.5
Resolution (Å)	50–2.40 (2.47–2.40)[a]
R_{sym}	0.10 (0.49, 0.77)[b]
$I / \sigma I$	15.3 (2.3, 1.7)[b]
Completeness (%)	88.7 (99.2, 36.2)[b]
Redundancy	4.7 (4.7, 4.2)[b]
Refinement	
Resolution (Å)	50–2.4
No. reflections	42,181
R_{work} / R_{free}	18.8 / 23.8 (26.5 / 35.2)[a]
No. atoms	6,717
Protein	6,259
Carbohydrate	177
Water	250
Ligands (PO$_4^-$, K$^+$)	31
B-factors (Å²)	
Protein	53
Carbohydrate	90
Water	54
Ligands (PO$_4^-$, K$^+$)	103
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.895

One crystal was used to collect the dataset.
[a]Values in parentheses are for the highest-resolution shell. [b]Because of anisotropic diffraction, data were integrated using an ellipsoidal mask with resolution cutoffs of 2.4 Å along the *a* and *b* axes and 2.7 Å along the *c* axis. Values in parentheses are given for the 2.85–2.7-Å resolution shell and for the highest-resolution shell, respectively.
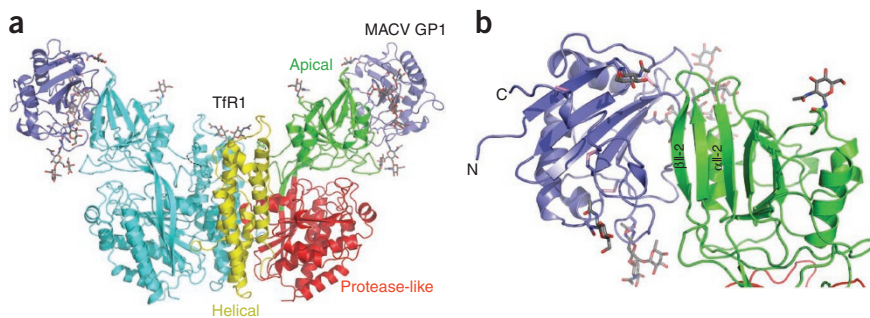
Figure 1 Structure of the MACV GP1–human TfR1 complex. (**a**) The butterfly-shaped TfR1 dimer is shown oriented with the cell surface at the bottom. One TfR1 protomer is colored according to the TfR1 domains: the apical domain is green, the protease-like domain is red and the helical domain is yellow. The other protomer is in cyan. MACV GP1, shown in purple, interacts with the TfR1 apical domain. The N-linked glycans are in gray. (**b**) Enlargement of the TfR1 apical domain with bound MACV GP1. Four N-linked glycans are on surfaces of the viral glycoprotein not involved in binding to TfR1. The N and C termini of MACV GP1 and the αII-2 helix and βII-2 strand of TfR1 are labeled.

Motif 1. Tyr211 in the apical domain of TfR1 is a critical determinant of MACV, JUNV and GTOV cell entry[18]. This residue, conserved in all orthologs of TfR1 that support the entry of New World arenaviruses, is a central feature of the GP1-TfR1 interface (**Fig. 3a**). Tyr211 is found within the hTfR1 βII-2 strand and fits snugly in a pocket on the viral glycoprotein. The hydroxyl group of TfR1 Tyr211 is hydrogen bonded with MACV GP1 Ser113, and its aromatic ring is surrounded by the side chains of MACV GP1 Arg111, Ile115 and Val117. The side chain of MACV GP1 Arg111 reaches past the aromatic ring to form a hydrogen bond with the main chain carbonyl of TfR1 Val210.

Motif 2. Asn348 participates in a well-defined hydrogen bonding network at the GP1-receptor interface (**Fig. 3b**). Asn348 caps the C terminus of αII-2 with a hydrogen bond to the backbone carbonyl of Lys344. It participates in an extensive hydrogen bonding network with MACV GP1 residues Asp114 and Ser116, and TfR1 residues Ser370 and Lys371. Mutation of TfR1 Asn348 to lysine interferes with the entry of

MACV, GTOV and in some contexts JUNV into cells expressing the altered receptor[18].

Motif 3. Val210 participates in hydrophobic interactions with the side chains of Phe226 and Tyr228 on MACV GP1 (**Fig. 3c**). These MACV GP1 residues both lie within loop 10, the disulfide-linked GP1 insert unique to MACV. The hydroxyl group of MACV GP1 Tyr228 hydrogen-bonds with the main chain amide of TfR1 Val210, and in addition to a side chain contact with Val210, MACV GP1 Phe226 is in van der Waals contact with receptor side chains of Ile201 and Leu212. MACV GP1 Pro223, also in loop 10, is in van der Waals contact with the δ-carbon of TfR1 Leu212. Mutation of Val210 in hTfR1 to glycine allows the altered receptor to support entry of Tacaribe virus (TCRV), a nonpathogenic New World arenavirus that is closely related to MACV and JUNV but does not otherwise use hTfR1 (ref. 23). We have not found a direct explanation for the effect of the V210G mutation on hTfR1 usage by TCRV. TfR1 Val210 is well anchored within a β-strand; substitution with a glycine may nonetheless allow sufficient local flexibility to compensate for some unfavorable aspect of the TCRV-hTfR1 interface.

Motif 4. Lys344, in the TfR1 αII-2 helix, packs against MACV GP1 Met119 and participates in a hydrogen bonding network that includes MACV GP1 Asp123 and Lys169 as well as TfR1 Glu294 (**Fig. 3d**). TfR1 Glu343, which precedes Lys344 in the αII-2 helix, accepts a hydrogen bond from the phenolic hydroxyl of MACV GP1 Tyr122. The Tyr122 ring also contacts TfR1 Ala293, Glu294 and Ala340, and Glu294 in turn has a salt bridge with GP1 Lys169. There is an additional hydrogen bond between the side chain of MACV GP1 Glu171 and the backbone amide of TfR1 Asn215. The contacts of motif 4 thus extend to the TfR1 loops between βII-2 and βII-3 and between βII-6 and βII-7 at the base of the apical domain.
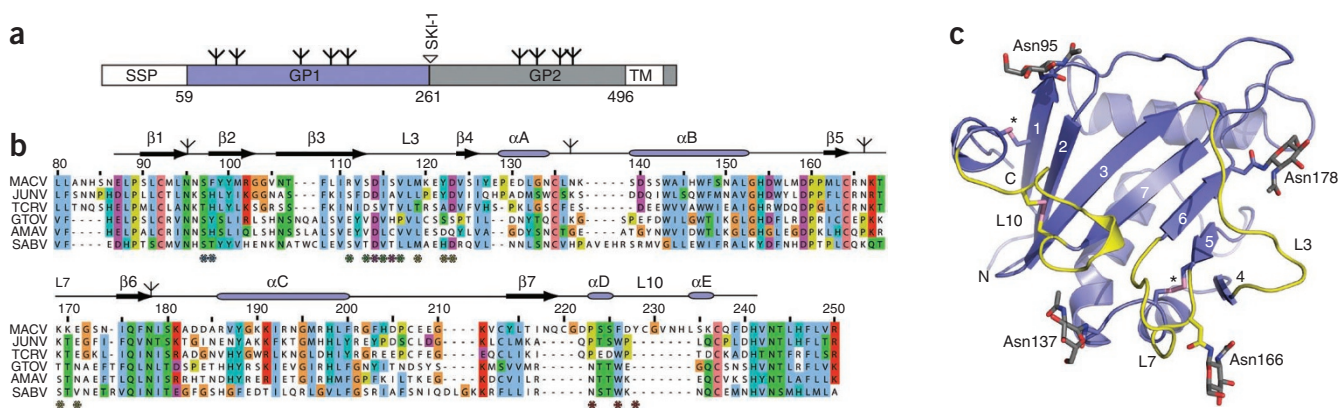

Figure 2 MACV GP1. (**a**) Domain organization of the MACV envelope glycoprotein (GP). GP is processed during its maturation into three noncovalently associated subunits: the stable signal peptide (SSP), GP1 and GP2. Sites of potential N-linked glycosylation are shown as tree diagrams. The triangle indicates the subtilisin/kexin isozyme 1 (SKI-1) cleavage site. TM, transmembrane segment. (**b**) Sequence alignment of MACV GP1 residues 80–250 with the GP1 proteins of NW clade B arenaviruses JUNV, TCRV, GTOV, AMAV and SABV. Secondary structure elements are shown for MACV GP1. Residues are colored according to the ClustalX color scheme. GP1 residues that interact with hTfR1 are labeled with asterisks colored according to the interaction motifs shown in **Figure 3**; green for motif 1, magenta for motif 2, red for motif 3, yellow for motif 4 and blue for motif 5. (**c**) Ribbon diagram of MACV GP1. The first N-acetylglucosamine is shown for glycosylation sites at Asn95, Asn137, Asn166 and Asn178. The four disulfide bonds that rigidify the GP1 core are in pink. The conserved Cys92-Cys237 and Cys135-Cys164 disulfide bridges are labeled with asterisks. The three loops that protrude from the GP1 core to contact TfR1 (L3, L7 and L10) are in yellow. Note that L10 is much shorter in the other viruses and will not contact TfR1.
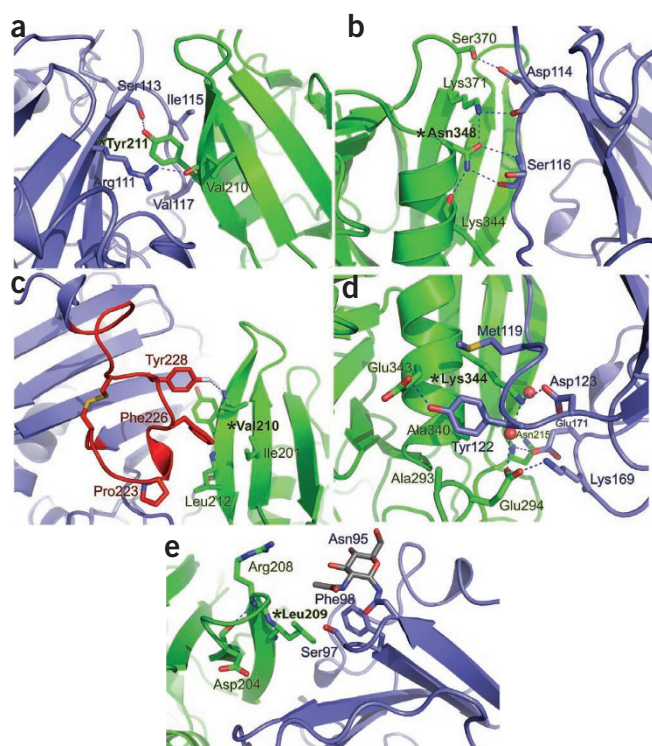
Figure 3 Five interaction motifs in the MACV GP1–human TfR1 interface. (**a**) Motif 1, centered on Tyr211 on the βII-2 strand in the receptor. (**b**) Motif 2, centered on TfR1 Asn348, which caps the C terminus of αII-2 helix. (**c**) Motif 3, centered on TfR1 Val210 in the βII-2 strand of the receptor, on the side opposite Tyr211. A segment unique to MACV GP1 is in red, with the Cys220–Cys229 disulfide bond in yellow. (**d**) Motif 4, at the lateral base of the TfR1 apical domain. An extensive network of polar interactions radiates from TfR1 Lys344. (**e**) Motif 5 at the lateral tip of the apical domain, centered on TfR1 Leu209. TfR1 residues Asp204 and Arg208, also shown, have been implicated previously in the host specificity of NW arenaviruses[18,23]. Asn95 is a conserved glycosylation site in New World arenavirus GP1s. The anchor residues for each motif are labeled with an asterisk. Hydrogen bonds are indicated as dashed lines.

Motif 5. Leu209 of TfR1 is in van der Waals contact with MACV GP1 Phe98 and the β-carbon of MACV GP1 Ser97 (**Fig. 3e**). Two nearby TfR1 residues, Arg208 and Asp204, have been implicated in the host specificity of New World arenaviruses. TfR1 Asp204 is replaced by an N-linked glycosylation site in the rodent host-species TfR1 orthologs (**Supplementary Fig. 3**), and removal of this glycan enhances the entry of MACV, JUNV and GTOV on the altered receptor[18]. Arg208, at the tip of the apical domain, is the first residue in the TfR1 βII-2 strand. In addition to the V210G mutation described above (motif 3), the R208G mutation, when introduced into hTfR1, also allows TCRV to use the altered receptor[23]. This change probably increases flexibility or adaptability of the loop between βII-1 and βII-2 of hTfR1.

Conservation of the GP1–TfR1 interface

We have examined the residue sequences and specificities of known New World arenavirus GP1s in view of the interface analysis just described. The sequences of MACV, JUNV, GTOV and SABV GP1 diverge substantially—based on the alignment shown in **Figure 2b**, the pairwise identities of MACV GP1 residues 87–242 with the corresponding GP1 sequences of JUNV, GTOV and SABV are 46%, 27% and 25%, respectively—but hTfR1 is a cell-entry receptor for all three viruses[17,19]. TCRV (41% identify with MACV GP1) can use hTfR1 when either Arg208 or Val210 in the hTfR1 βII-2 strand is mutated to glycine, suggesting that modest changes in its GP1 protein might also allow it to bind human TfR1 (ref. 23). Amapari virus (AMAV, 29% GP1 sequence identity with MACV) can use hTfR1 when four mutations are made to the βII-1–βII-2 loop and βII-2 strand in the receptor[23]. Thus, all six GP1 receptor-binding domains appear to be complementary to the same surface. Location of conserved residues in GP1 justifies the further assumption that, despite considerable sequence divergence, all New World arenavirus GP1s have the same framework structure.

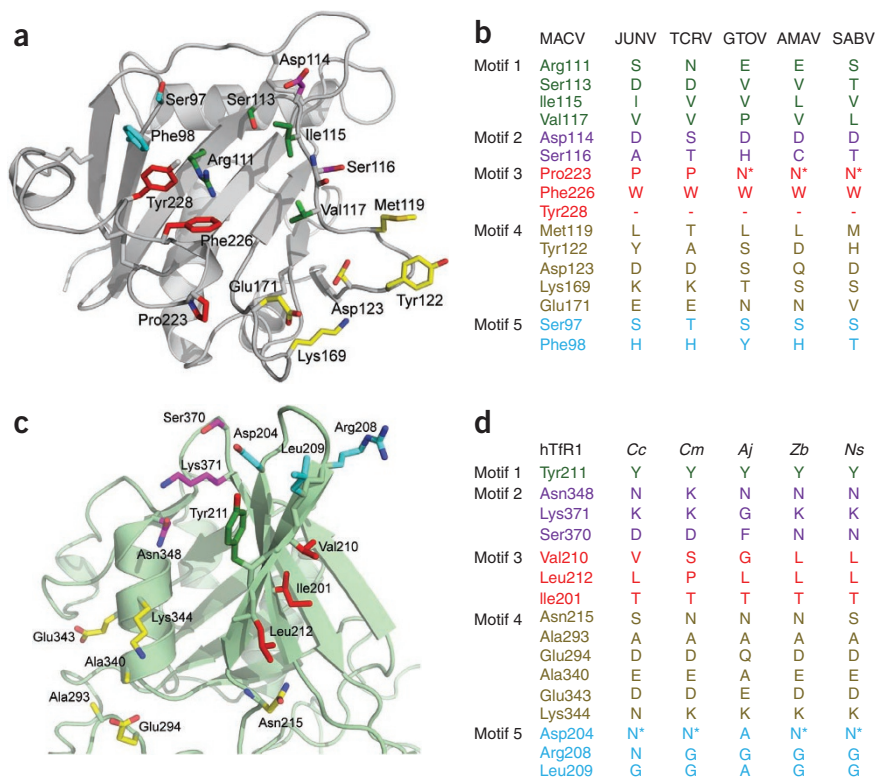MACV GP1 has two long-range disulfide bonds (one between Cys92 and Cys237, bounding the entire domain, and the other between Cys135 and Cys164, bridging across the hydrophobic core) and two short-range disulfide bonds (from Cys207 to Cys213 and from Cys220 to Cys229, stabilizing local loops). The alignment in **Figure 2b** shows that the two long-range disulfide bonds are conserved and that the short-range disulfides are unique to MACV. The alignment predicts that GTOV GP1 will have an additional long-range disulfide between Cys120 and Cys164 (MACV numbering) and that AMAV GP1 will have a long-range disulfide between Cys116 and Cys213 (MACV numbering). Both disulfide pairings are fully compatible with a conserved fold. SABV GP1 has a short insertion following residue 107 (MACV numbering) with an unpaired cysteine, which projects away from the TfR1 interface and probably bridges to a cysteine elsewhere in the viral glycoprotein.

GP1 residues contacting TfR1

Figure 4a shows the viral side of the GP1 interface, with GP1 residues grouped according to the five interaction motifs. In viral contributions to interaction motif 1, the hydrogen bond between MACV GP1 Ser113 and the phenolic hydroxyl group of TfR1 Tyr211 can also be made by the threonine at that position in SABV and by the aspartate in JUNV and TCRV GP1 (**Fig. 4b**). GTOV and AMAV, however, have a hydrophobic residue at the equivalent of position 113; a glutamate nearby (replacing MACV GP1 Arg111) might accept the hydrogen bond instead. The residue at GP1 position 115, which lines the pocket occupied by TfR1 Tyr211, is isoleucine in MACV and JUNV, leucine in AMAV and valine in TCRV, GTOV and SABV. Residue 117, which contributes the other hydrophobic side chain lining the GP1 pocket for TfR1 Tyr211, is valine in MACV, JUNV, TCRV and AMAV, leucine in SABV and proline in GTOV. MACV GP1 Arg111, which makes prominent contacts with the TfR1 βII-2 strand, is not conserved in any of the other viruses. In motif 2, TfR1 Asn348 has reciprocal hydrogen bond interactions with the main chain amide and carbonyl of MACV GP1 Ser116; the side chain of this residue does not participate, nor is it conserved in any of the GPs. Interaction motif 3 is unique to MACV, as it involves the insert in loop 10. In motif 4, three of five MACV GP1 residues that contact TfR1 are conserved in the other clade B1 viruses TCRV and JUNV, and four of five are conserved between MACV and JUNV. However, substantial variability is seen in this motif when GTOV, AMAV and SABV are compared with the clade B1 viruses. In motif 5, Phe98 is replaced by a more polar residue in the other viruses, as might be expected, as absence of loop 10 would leave this phenylalanine substantially more exposed.

TfR1 residues contacting GP1

The New World hemorrhagic fever viruses vary in their capacity to infect cultured cells expressing the TfR1 orthologs from the natural hosts[18,23] (**Table 1**). For example, MACV can use its host-species TfR1 ortholog (*Cc*TfR1, from *Calomys callosus*) as a receptor, but

Figure 4 Conservation of the GP1-TfR1 interface. (**a**) Ribbon diagram of MACV GP1, as seen from the apical domain of TfR1 in the complex. GP1 residues that contact TfR1 are colored according to the interaction motifs shown in **Figure 3**, using the color scheme outlined in the caption to **Figure 2b**. (**b**) List of MACV GP1 residues that contact TfR1, colored according to the five interaction motifs. Residues found in analogous positions for the New World clade B arenaviruses JUNV, TCRV, GTOV, AMAV and SABV (based on the sequence alignment shown in **Fig. 2b**) are also shown. (**c**) Side-view ribbon diagram of the TfR1 apical domain. TfR1 residues that contact MACV GP1 are colored according to the five interaction motifs as in **a**. (**d**) List of hTfR1 residues that contact MACV GP1, colored according to the five interaction motifs. Residues found in analogous positions in *C. callosus* (MACV host), *C. musculinus* (JUNV host), *A. jamaicensis* (TCRV host), *Z. brevicauda* (GTOV host) and *N. spinosus* (AMAV host) TfR1 are listed based on sequence alignment (see **Supplementary Fig. 3**).

b

	MACV	JUNV	TCRV	GTOV	AMAV	SABV
Motif 1	Arg111	S	N	E	E	S
	Ser113	D	D	V	V	T
	Ile115	I	V	V	L	V
	Val117	V	V	P	V	L
Motif 2	Asp114	D	S	D	D	D
	Ser116	A	T	H	C	T
Motif 3	Pro223	P	P	N*	N*	N*
	Phe226	W	W	W	W	W
	Tyr228	-	-	-	-	-
Motif 4	Met119	L	T	L	L	M
	Tyr122	Y	A	S	D	H
	Asp123	D	D	S	Q	D
	Lys169	K	K	T	S	S
	Glu171	E	E	N	N	V
Motif 5	Ser97	S	T	S	S	S
	Phe98	H	H	Y	H	T

d

	hTfR1	*Cc*	*Cm*	*Aj*	*Zb*	*Ns*
Motif 1	Tyr211	Y	Y	Y	Y	Y
Motif 2	Asn348	N	K	N	N	N
	Lys371	K	K	G	K	K
	Ser370	D	D	F	N	N
Motif 3	Val210	V	S	G	L	L
	Leu212	L	P	L	L	L
	Ile201	T	T	T	T	T
Motif 4	Asn215	S	N	N	N	S
	Ala293	A	A	A	A	A
	Glu294	D	D	Q	D	D
	Ala340	E	E	A	E	E
	Glu343	D	D	E	D	D
	Lys344	N	K	K	K	K
Motif 5	Asp204	N*	N*	A	N*	N*
	Arg208	N	G	G	G	G
	Leu209	G	G	A	G	G

not that of the JUNV host species (*Cm*TfR1, from *Calomys musculinus*)[18]. GTOV can use only the TfR1 ortholog of its host (*Zb*TfR1, from *Zygodontomys brevicauda*) and that of the closely related New World arenavirus AMAV (*Ns*TfR1, from *Neacomys spinosus*)[18,23]. The restricted receptor usage of each virus clearly reflects a relatively long period of virus-host co-adaptation[24].

Figure 4c shows the receptor side of the GP1-TfR1 interface. TfR1 Tyr211, the central feature of motif 1 (**Fig. 3a**), is conserved in all the host-species receptors (**Fig. 4d**). TfR1 Asn348, which anchors the second interaction motif (**Fig. 3b**), is likewise conserved, except in *Cm*TfR1 (JUNV host), which has a lysine at position 348. MACV cannot use *Cm*TfR1 as a receptor, and the N348K substitution in the context of hTfR1 consistently disrupts the entry of MACV and GTOV into cells expressing this altered receptor[18]. The lysine at 348 would disrupt the hydrogen bond network that includes the C cap on TfR1 αII-2 and the main chain hydrogen bonds with MACV GP1 Ser116 (**Fig. 3b**). Presumably, JUNV, in adapting to the receptor of its native host, has evolved to accommodate Lys348 in binding to *Cm*TfR1. In motif 3, the interaction unique to MACV, most of the receptors present a hydrophobic surface (valine or leucine at position 210 and leucine at position 212: **Fig. 3c**). The exceptions are *Cm*TfR1 and *Aj*TfR1 (from *Artibeus jamaicensis*). The former has a serine at position 210 and a proline at position 212; the latter, a glycine at position 210. Loop 10 of MACV GP1 engages this hydrophobic surface, with MACV GP1 Phe226 inserted between the side chains of TfR1 residues 210 and 212. The details of the contact suggest that Cγ of proline at position 212 would clash with the Phe226 side chain, consistent with the failure of MACV efficiently to infect cells expressing *Cm*TfR1 (ref. 18).

The anchoring residue of motif 4, TfR1 Lys344, is conserved in all receptors except that of the MACV host (*Cc*TfR1); its substitution by asparagine could lead to plausible rearrangements of the hydrogen bond and salt-bridge network at the contact (**Fig. 3d**). Another difference in motif 4 is TfR1 residue 294, which is aspartate in the host receptor for MACV (and also those for JUNV, GTOV and AMAV) but glutamate in hTfR1. An aspartate side chain at

this position cannot reach the ε-amino group of GP1 Lys169 (conserved in the clade B1 viruses), but the salt bridge can be replaced by one from TfR1 residue 340, which is glutamate rather than alanine in precisely those receptors that have the aspartate substitution at position 294.

The preceding analysis of contact motifs at the extensive GP1–apical domain interface illustrates that the sometimes substantial residue-sequence differences among the various viruses and receptors are nonetheless consistent with an overall retention of local conformation within each motif, as key contacts have conserved characteristics. Moreover, the conservation of contacts among the viruses and their native host-species receptors at the base of the apical domain (motif 4), and the conserved packing of GP1 residues around TfR1 Tyr211 some 20 Å away near the tip of the apical domain (motif 1), together indicate that GP1 will dock against TfR1 in essentially the same orientation and position in any productive virus-receptor pairing. The general consistency of the analysis indicates that the interface between alternate pairings of GP1 proteins and receptors varies only in detail, not in the overall mode in which one partner docks against the other.

Recognition of human TfR1
Among the TfR1 sequences we have compared, Arg208 is a unique feature of the human receptor. The corresponding position is occupied by glycine in all New World arenavirus host-species receptors examined here, with the exception of the MACV host (*Cc*TfR1), in which the position is occupied by an asparagine (**Fig. 4d**). The R208G mutation converts hTfR1 into a receptor for TCRV[23]. Because most of the viruses are probably best adapted to the absence of a large side chain at the 208 position in TfR1, we asked whether an R208G change in hTfR1 also affects the infectivity of MACV, JUNV and GTOV. We generated fusion proteins comprising the truncated GP1 subunits of MACV, JUNV or GTOV fused to the Fc portion of
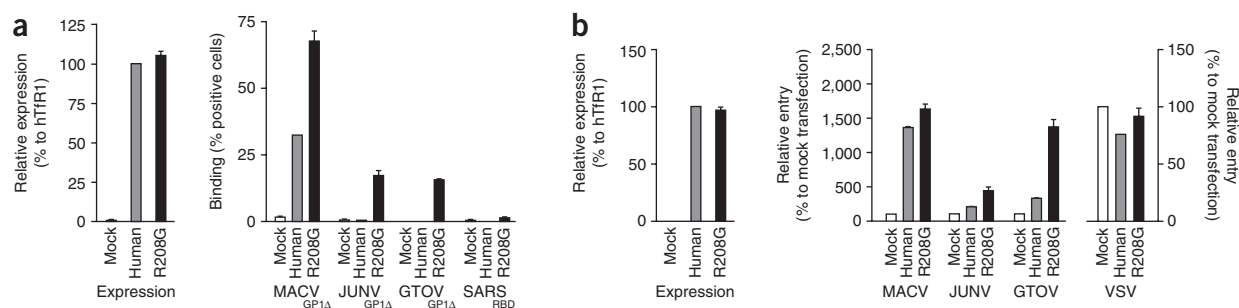
Figure 5 Importance of the tip of the apical domain in determining the role of human TfR1 as a receptor for pathogenic New World arenaviruses. (**a**) CHO cells were transfected with vector alone (mock) or plasmids encoding hTfR1 (human), or hTfR1 with the Arg208Gly mutation (R208G); 48 h later, the cells were treated with an anti-Flag antibody to measure cell surface expression of receptors (left), or with immunoglobulin (Ig) fusion proteins containing the truncated GP1 subunits (GP1Δ) of MACV, JUNV or GTOV or the receptor binding domain of the severe acute respiratory syndrome coronavirus (SARS-CoV) spike protein (right). Association of these proteins with cells was measured by flow cytometry. (**b**) CHO cells were transfected with the indicated plasmids, and cell surface expression of the receptors (left) determined 48 h later with an anti-Flag antibody as in (*a*). Aliquots of these cells were infected in parallel with MACV, JUNV, GTOV or vesicular stomatitis virus (VSV) pseudoviruses expressing GFP. Infection levels were assessed 48 h later by measuring GFP expression by flow cytometry. Error bars indicate s.d.

human IgG (GP1Δ-Ig)[17] and tested them for binding to CHO cells transfected with wild-type (WT) hTfR1 or hTfR1 R208G. MACV GP1Δ-Ig bound CHO cells expressing either WT hTfR1 or hTfR1 R208G but bound the latter more strongly (**Fig. 5a**). JUNV and GTOV GP1Δ-Ig bound only cells expressing hTfR1 R208G. Lack of detectable binding by JUNV and GTOV GP1Δ-Ig to cells expressing WT hTfR1 probably resulted from the low affinity of these proteins for the wild-type receptor. Consistent with these data, the R208G mutation modestly enhanced infectious entry of MACV pseudovirus but led to substantial increases in infection by JUNV and GTOV pseudoviruses (**Fig. 5b**). The tip of the TfR1 apical domain (motif 5) is thus a determinant of entry by both pathogenic and nonpathogenic New World clade B arenaviruses. The rodent receptors all have a simple, four-residue β-turn in the βII-1–βII-2 loop, whereas human TfR1 has a five-residue turn. The pathogenic viruses (MACV, JUNV and GTOV) have presumably adapted in their natural hosts to the shorter turn, which places a glycine at the 209 position instead of leucine. The substitution eliminates the hydrophobic contact we see between hTfR1 Leu209 and MACV GP1 Phe98 (**Fig. 3e**). The alternative interactions that we have suggested might be present in different virus-host pairs vary in character and position, and we therefore cannot readily predict from GP1 sequences alone how many residue changes might be needed to make any particular nonpathogenic New World clade B arenavirus infectious in humans without compromising ability to propagate in its natural host.

DISCUSSION

The structure we have determined shows that MACV GP1 contacts the apical domain of TfR1. Although present in all mammalian TfR1s, the apical domain does not have a known physiological role: neither transferrin nor the hereditary hemochromatosis associated protein, HFE, interact with it[22,25]. As New World arenavirus GPs engage TfR1 in a manner that does not seem to interfere with binding of its known physiological ligands[17], antibodies that target the TfR1 apical domain might in principle be useful as broadly-acting therapeutic agents for acute treatment of New World hemorrhagic fevers.

Our structure, when compared with the structure of unliganded MACV GP1 (ref. 26), shows that no significant conformational change occurs upon receptor binding (Cα r.m.s. deviation = 0.55). Receptor binding is therefore unlikely to prime GP for pH-induced membrane fusion before virions are internalized into endosomes as seen, for

example, with the avian leukosis virus[27,28]. The events during New World arenavirus cell entry may be more similar to those during entry of influenza viruses into cells, in which hemagglutinin attachment to a cell-surface receptor (sialic acid) serves primarily to traffick virions to acidified endosomes rather than to initiate a series of conformational rearrangements[14].

The structure of MACV GP1 bound to human TfR1 reveals that New World clade B arenaviruses contact TfR1 on a surface with substantial sequence variability among viruses. This sequence variability, which presumably reflects the independent evolution of individual viruses in their various host species[24], is best illustrated by the insert in MACV loop 10, which forms extensive receptor contacts yet is much shorter and lacks these contacts in the other viruses (**Fig. 3b**). Nonetheless, a number of similarities among the GP1 receptor-binding domains strongly support the assumption of a shared fold within members of the family. Conservation of key receptor residues in several of the interaction motifs indicates conservation of local features of the interface as well as of the overall docking of the two partners. The structure of the complex may therefore prove useful in predicting whether a novel New World arenavirus is likely to use TfR1. For example, extending the analysis shown in **Figure 4** to the recently isolated New World hemorrhagic fever arenavirus Chapare suggests that this virus will indeed use TfR1 as its receptor[8] (**Supplementary Fig. 4**).

The loop between the TfR1 βII-1 and βII-2 strands appears to be a key determinant of whether any particular GP1 can bind with adequate affinity to a TfR1 ortholog[18,23]. That loop is longer by one residue in human TfR1 than in its ortholog in any of the New World arenavirus native hosts, except for that of TCRV. Moreover, at the end of that loop, Arg208 in the human receptor becomes glycine in the host-species receptors (with the exception of the MACV host, where it is an asparagine), and the R208G mutation enhances infectivity of JUNV and GTOV and allows infection by TCRV. The large arginine residue thus presents a barrier to use of human TfR1 by both nonpathogenic[23] and pathogenic New World clade B arenaviruses (**Fig. 5**). The New World clade B arenaviruses appear to be adapted best to their host-species receptors. We suggest that those infecting humans accidentally acquired the capacity to bind human TfR1 during coevolution with the receptor in their natural hosts. The structure of MACV GP1 bound to hTfR1 thus clarifies the complex structural basis for the zoonotic transmission of this important class of emerging human pathogens.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/nsmb/.

Accession codes. The coordinates for the structure of MACV GP1 bound to hTfR1 have been deposited in the Protein Data Bank with accession code 3KAS.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

AUTHOR CONTRIBUTIONS

J.A. designed and performed the experiments, analyzed the data and wrote the paper; K.D.C. assisted with data collection, molecular replacement, re-interpretation of the unliganded TfR1 structures and edited the paper; M.F. and H.C. assisted with data analysis and interpretation and edited the paper; S.C.H. helped design experiments, advised on model building and interpretation and participated in writing and editing the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nsmb/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Oldstone, M.B. Arenaviruses. I. The epidemiology molecular and cell biology of arenaviruses. Introduction. *Curr. Top. Microbiol. Immunol.* **262**, V–XII (2002).
2. Bowen, M.D., Peters, C.J. & Nichol, S.T. The phylogeny of New World (Tacaribe complex) arenaviruses. *Virology* **219**, 285–290 (1996).
3. Clegg, J.C. Molecular phylogeny of the arenaviruses. *Curr. Top. Microbiol. Immunol.* **262**, 1–24 (2002).
4. Emonet, S., Lemasson, J.J., Gonzalez, J.P., de Lamballerie, X. & Charrel, R.N. Phylogeny and evolution of old world arenaviruses. *Virology* **350**, 251–257 (2006).
5. Charrel, R.N. & de Lamballerie, X. Arenaviruses other than Lassa virus. *Antiviral Res.* **57**, 89–100 (2003).
6. Lisieux, T. *et al.* New arenavirus isolated in Brazil. *Lancet* **343**, 391–392 (1994).
7. Tesh, R.B., Jahrling, P.B., Salas, R. & Shope, R.E. Description of Guanarito virus (*Arenaviridae*: Arenavirus), the etiologic agent of Venezuelan hemorrhagic fever. *Am. J. Trop. Med. Hyg.* **50**, 452–459 (1994).
8. Delgado, S. *et al.* Chapare virus, a newly discovered arenavirus isolated from a fatal hemorrhagic fever case in Bolivia. *PLoS Pathog.* **4**, e1000047 (2008).
9. Buchmeier, M.J. Arenaviruses: protein structure and function. *Curr. Top. Microbiol. Immunol.* **262**, 159–173 (2002).
10. Eichler, R. *et al.* Identification of Lassa virus glycoprotein signal peptide as a trans-acting maturation factor. *EMBO Rep.* **4**, 1084–1088 (2003).
11. Saunders, A.A. *et al.* Mapping the landscape of the lymphocytic choriomeningitis virus stable signal peptide reveals novel functional domains. *J. Virol.* **81**, 5649–5657 (2007).
12. York, J. & Nunberg, J.H. Role of the stable signal peptide of Junin arenavirus envelope glycoprotein in pH-dependent membrane fusion. *J. Virol.* **80**, 7775–7780 (2006).
13. York, J. & Nunberg, J.H. Distinct requirements for signal peptidase processing and function in the stable signal peptide subunit of the Junin virus envelope glycoprotein. *Virology* **359**, 72–81 (2007).
14. Harrison, S.C. Viral membrane fusion. *Nat. Struct. Mol. Biol.* **15**, 690–698 (2008).
15. Rojek, J.M. & Kunz, S. Cell entry by human pathogenic arenaviruses. *Cell. Microbiol.* **10**, 828–835 (2008).
16. York, J., Agnihothram, S.S., Romanowski, V. & Nunberg, J.H. Genetic analysis of heptad-repeat regions in the G2 fusion subunit of the Junin arenavirus envelope glycoprotein. *Virology* **343**, 267–274 (2005).
17. Radoshitzky, S.R. *et al.* Transferrin receptor 1 is a cellular receptor for New World haemorrhagic fever arenaviruses. *Nature* **446**, 92–96 (2007).
18. Radoshitzky, S.R. *et al.* Receptor determinants of zoonotic transmission of New World hemorrhagic fever arenaviruses. *Proc. Natl. Acad. Sci. USA* **105**, 2664–2669 (2008).
19. Flanagan, M.L. *et al.* New World clade B arenaviruses can use transferrin receptor 1 (TfR1)-dependent and independent entry pathways, and glycoproteins from human pathogenic strains are associated with the use of TfR1. *J. Virol.* **82**, 938–948 (2008).
20. Aisen, P. Transferrin receptor 1. *Int. J. Biochem. Cell Biol.* **36**, 2137–2143 (2004).
21. Lawrence, C.M. *et al.* Crystal structure of the ectodomain of human transferrin receptor. *Science* **286**, 779–782 (1999).
22. Cheng, Y., Zak, O., Aisen, P., Harrison, S.C. & Walz, T. Structure of the human transferrin receptor-transferrin complex. *Cell* **116**, 565–576 (2004).
23. Abraham, J. *et al.* Host-species transferrin receptor 1 orthologs are cellular receptors for nonpathogenic new world clade B arenaviruses. *PLoS Pathog.* **5**, e1000358 (2009).
24. Bowen, M.D., Peters, C.J. & Nichol, S.T. Phylogenetic analysis of the *Arenaviridae*: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. *Mol. Phylogenet. Evol.* **8**, 301–316 (1997).
25. Bennett, M.J., Lebron, J.A. & Bjorkman, P.J. Crystal structure of the hereditary haemochromatosis protein HFE complexed with transferrin receptor. *Nature* **403**, 46–53 (2000).
26. Bowden, T.A. *et al.* Unusual molecular architecture of the machupo virus attachment glycoprotein. *J. Virol.* **83**, 8259–8265 (2009).
27. Mothes, W., Boerger, A.L., Narayan, S., Cunningham, J.M. & Young, J.A. Retroviral entry mediated by receptor priming and low pH triggering of an envelope glycoprotein. *Cell* **103**, 679–689 (2000).
28. Smith, J.G., Mothes, W., Blacklow, S.C. & Cunningham, J.M. The mature avian leukosis virus subgroup A envelope glycoprotein is metastable, and refolding induced by the synergistic effects of receptor binding and low pH is coupled to infection. *J. Virol.* **78**, 1403–1410 (2004).

ONLINE METHODS

Expression, purification and crystallization of MACV GP1 and human TfR1. We expressed MACV GP1$_{79-258}$ with the addition of a hexahistidine tag at the C terminus in Sf9 insect cells and purified the protein by nickel affinity chromatography. We subjected the protein to limited proteolysis with elastase and further purified it by size-exclusion chromatography. We purified the soluble ectodomain of hTfR1 as previously described[21]. We concentrated both proteins separately in buffer containing 25 mM Tris, pH 7.5 and 150 mM NaCl and mixed them in equimolar ratio to a final concentration of 10 mg ml^{-1}. Crystals grew in 1.6 M NaKPO$_4$, pH 5.5, 150 mM NaCl. We flash froze crystals in mother liquor containing 20% (v/v) glycerol for data collection.

Data collection and structure determination. The MACV GP1–hTfR1 complex crystallized in the $C222_1$ space group, with unit cell dimensions a = 140 Å, b = 173 Å and c = 104 Å.

We collected X-ray diffraction data at 100 K at NE-CAT beamline ID-24C at the Advanced Photon Source (Argonne National Laboratory). Because of anisotropic diffraction, we integrated the data using an ellipsoidal mask with resolution cutoffs of 2.4 Å along the a and b axes and 2.7 Å along the c axis[29]. We processed the data using HKL2000 (http://www.hkl-xray.com) and determined the structure of the complex by molecular replacement with PHASER[30], using the hTfR1 monomer[21] (PDB 1CX8) as a search model. We built the model into density with Coot[31]. Iterative model building and refinement with PHENIX[32] yielded final R_{free} of 24% and R_{work} of 20% (**Table 2**); 96.8% of all residues were in the most favored region of Ramachandran space, and 99.9% (782/783 residues) were in allowed regions. We calculated buried surface areas and simulated-annealing omit maps (**Supplementary Fig. 5**) with CNS[33] and made figures with PyMol[34].

Sequence alignment. We used ClustalW[35] to align the New World arenaviral GP1 and TfR1 sequences and adjusted the positions of gaps in the sequence alignments by referring to the three-dimensional structures. We selected the following glycoprotein sequences: AMAV (strain BeAn 70563, GenBank accession number AF512834), Chapare virus (810419, NC_010562), GTOV (INH-95551, NC_005077), JUNV (MC2, D10072), MACV (Carvallo, NC_005078) and SABV (SPH114202, NC_006317). The sequence of TCRV GP (a gift from S. Kunz, University of Lausanne, Switzerland) is that of a laboratory-adapted strain, which has a deletion spanning residues 121–132 and substitutions at three residues (I134A, G418S and E458R) compared to GenBank NC_004293. We used the following TfR1 sequences: human TfR1 (NM_003234), *C. callosus* TfR1 (EU164540), *C. musculinus* TfR1 (EU164541), *Z. brevicauda* TfR1 (EU340259), *A. jamaicensis* TfR1 (FJ154605) and *N. spinosus* TfR1 (FJ154604). We made figures and did pairwise sequence identity calculations using Jalview[36,37].

Cells and plasmids. We maintained HEK 293T cells (human embryonic kidney, ATCC, CRL-11268) in DMEM and CHO epithelial cells (ATCC, CCL-61) in F12 medium in the presence of 10% (v/v) FBS. Plasmids encoding hTfR1 and hTfR1 R208G, with the addition of a C-terminal Flag tag, have been described previously[23], as have MACV (Carvallo strain), JUNV (MC2) and GTOV (INH-95551) GP-expressor plasmids[17].

GP1Δ-Ig binding assays. We generated MACV GP1Δ-Ig, JUNV GP1Δ-Ig, and GTOV GP1Δ-Ig by PCR amplification of MACV GP1 residues 87–258, JUNV GP1 residues 87–241 and GTOV GP1 residues 79–228. We cloned these fragments into a previously described pcDM8-based plasmid containing the CD5 signal sequence and the Fc region of human IgG1 (ref. 38). The SARS-CoV RBD Ig fusion protein has been described previously[38]. For Ig fusion protein production, we transfected HEK 293T cells with the appropriate plasmid and harvested the supernatant at 48 h after transfection. For binding assays, we transfected CHO cells with hTfR1 or hTfR1 R208G. After 48 h, we detached the cells with 1 mM EDTA in PBS and incubated them in 100 μl of PBS containing 2% goat serum and 0.5 μg of an anti-Flag M2 monoclonal antibody (Sigma) or supernatant containing 0.5 μg of the indicated Ig fusion protein. We then stained the cells with anti-mouse IgG or with anti-human IgG (Fc-specific) antibodies conjugated with phycoerythrin (Jackson Immuno Laboratories).

Pseudovirus infection. We generated recombinant retroviruses by transfecting HEK 293T cells in a 1:1:1 ratio of plasmids encoding the respective arenaviral GP, the MLV *gag* and *pol* genes and the pQCXIX retroviral vector (BD Biosciences) expressing GFP, as previously described[17]. We harvested cell supernatants 24 h after transfection and filtered them through a 0.45-μm filter. We transfected CHO cells with plasmids expressing WT hTfR1 or hTfR1 R208G and replated the cells at 24 h for flow cytometry and pseudovirus infection. At 48 h after transfection, we assessed TfR1 expression levels by flow cytometry using an anti-Flag M2 antibody. In parallel, we infected the cells with the pseudoviruses for 1 h. We detached the cells with trypsin 48 h after transfection and measured the GFP-expression levels by flow cytometry. We normalized GFP-expression levels to that of mock-transfected cells.

29. Strong, M. *et al.* Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* **103**, 8060–8065 (2006).
30. McCoy, A.J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
31. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
32. Adams, P.D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954 (2002).
33. Brunger, A.T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2**, 2728–2733 (2007).
34. DeLano, W.L. *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, California, USA, 2002).
35. Thompson, J.D., Higgins, D.G. & Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
36. Clamp, M., Cuff, J., Searle, S.M. & Barton, G.J. The Jalview Java alignment editor. *Bioinformatics* **20**, 426–427 (2004).
37. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. & Barton, G.J. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
38. Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).