CD4 receptor diversity in chimpanzees protects against SIV infection

Citation for published version:
https://doi.org/10.1073/pnas.1821197116

Digital Object Identifier (DOI):
10.1073/pnas.1821197116

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Proceedings of the National Academy of Sciences

Publisher Rights Statement:
This open access article is distributed under Creative Commons Attribution-NonCommercial NoDerivatives License 4.0 (CC BY-NC-ND). https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
CD4 receptor diversity in chimpanzees protects against SIV infection


aDepartment of Medicine, University of Pennsylvania, Philadelphia, PA 19104; bDepartment of Microbiology, University of Pennsylvania, Philadelphia, PA 19104; cVaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892; dInstitute of Evolutionary Biology, University of Edinburgh, EH9 3FL Edinburgh, United Kingdom; eDepartment of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294; fDepartment of Medicine, University of California at Berkeley, Berkeley, CA 94720; gDepartment of Evolutionary Anthropology, Duke University, Durham, NC 27708

Contributed by Beatrice H. Hahn, December 26, 2018 (sent for review December 13, 2018; reviewed by Facundo Batista and Jeremy Luban)

Human and simian immunodeficiency viruses (HIV/SIVs) use CD4 as the primary receptor to enter target cells. Here, we show that the chimpanzee CD4 is highly polymorphic, with nine coding variants present in wild populations, and that this diversity interferes with SIV envelope (Env)-CD4 interactions. Testing the replication fitness of SIVcpz strains in CD4+ T cells from captive chimpanzees, we found that certain viruses were unable to infect cells from certain hosts. These differences were recapitulated in CD4 transfection assays, which revealed a strong association between CD4 genotype and SIVcpz infection phenotypes. The most striking differences were observed for three substitutions (Q25R, Q40R, and P68T), with P68T generating a second N-linked glycosylation site (N66) in addition to an invariant N32 encoded by all chimpanzee CD4 alleles. In silico modeling and site-directed mutagenesis identified charged residues at the CD4-Env interface and clashes between CD4- and Env-encoded glycans as mechanisms of inhibition. CD4 polymorphisms also reduced Env-mediated cell entry of monkey SIVs, which was dependent on at least one D1 domain glycan. CD4 allele frequencies varied among at least one D1 domain glycan. CD4 allele frequencies varied among chimpanzees, protecting chimpanzees against SIV infection. Although some SIVcpz strains have adapted to utilize these variants, CD4 diversity is maintained, providing chimpanzees against infection with SIVcpz and other SIVs to which they are exposed.

Significance


Reviewers: F.B., Massachusetts General Hospital; and J.L., University of Massachusetts Medical School.

The authors declare no conflict of interest.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: Chimpanzee mitochondrial D-loop, CD4 coding, and CD4 exons 2 and 3 sequences have been deposited in the GenBank database (accession nos. MK178722-MK178852, MK208839-MK208863, and MK208839-MK208845, respectively). Analysis code is archived on Zenodo (doi: https://zenodo.org/record/2527032#.XDAqz9IzbIU).

Human and simian immunodeficiency viruses (HIV/SIVs) use CD4 as the primary receptor to enter target cells. Here, we show that the chimpanzee CD4 is highly polymorphic, with nine coding variants present in wild populations, and that this diversity interferes with SIV envelope (Env)-CD4 interactions. Testing the replication fitness of SIVcpz strains in CD4+ T cells from captive chimpanzees, we found that certain viruses were unable to infect cells from certain hosts. These differences were recapitulated in CD4 transfection assays, which revealed a strong association between CD4 genotype and SIVcpz infection phenotypes. The most striking differences were observed for three substitutions (Q25R, Q40R, and P68T), with P68T generating a second N-linked glycosylation site (N66) in addition to an invariant N32 encoded by all chimpanzee CD4 alleles. In silico modeling and site-directed mutagenesis identified charged residues at the CD4-Env interface and clashes between CD4- and Env-encoded glycans as mechanisms of inhibition. CD4 polymorphisms also reduced Env-mediated cell entry of monkey SIVs, which was dependent on at least one D1 domain glycan. CD4 allele frequencies varied among at least one D1 domain glycan. CD4 allele frequencies varied among chimpanzees, protecting chimpanzees against SIV infection. Although some SIVcpz strains have adapted to utilize these variants, CD4 diversity is maintained, protecting chimpanzees against infection with SIVcpz and other SIVs to which they are exposed. CD4 | SIV | chimpanzee | envelope glycoprotein | glycan restriction
TRIM5, and SAMHD1, have suggested that SIVs have been infecting primates for millions of years (6–8), which is consistent with their complex evolutionary history characterized by virus-host coevolution, cross-species transmission, recombination, and lineage extinction (9–13). SIVs can be highly pathogenic in both primate and human hosts (14, 15), with SIVcpz from chimpanzee having generated HIV type 1 (HIV-1), the cause of the AIDS pandemic (16). Primate hosts with more longstanding SIV infections, such as African green monkeys (Cercopithecus aethiops) and sooty mangabeys (Cercopithecus ascanius), have evolved mechanisms that shield them from the pathogenic effects of their SIV strains (15, 17–19). However, such protective barriers are absent when SIVs invade new, nonadapted hosts.

Chimpanzees (P. troglodytes) are comprised of four geographically distinct subspecies, including western (Pan troglodytes verus), Nigeria-Cameroonian (Pan troglodytes elioti), central (Pan troglodytes troglodytes), and eastern (Pan troglodytes schweinfurthii) chimpanzees (SI Appendix, Fig. S1). Only the central and eastern subspecies are SIVcpz-infected (16, 20–22), suggesting that this virus was introduced after the divergence and geographic separation of the other two subspecies. Analysis of the SIVcpz genome has shown that chimpanzees are not an obligate host for SIV, which implies that the virus was brought into them by hunting. This divergence in SIVcpz prevalence may be explained by the introduction of chimpanzees to West Africa by 400,000 years ago (23). The central and eastern chimpanzees that are SIVcpz-infected are predominantly housed in US primate centers and infected their CD4+ T cells with a panel of eight chimpanzee viruses representing both SIVcpz strains (MT145, EK505, MB897, LB715, and GAB2) and SIVcpZPs (BF1167, TAN2, and TAN13) strains (21, 44–47). SIVcpZPs were transfected, normalized based on infectivity in a permissive cell line (TZM-bi), and used to infect CD4+ T cells at a multiplicity of infection of 0.1 (Fig. 1). As observed previously (21, 44, 47), all SIVcpZP strains replicated efficiently in human CD4+ T cells (Fig. 1A). However, their ability to establish a productive infection in chimpanzee CD4+ T cells varied considerably. Of the 28 chimpanzees tested, only 9 supported the replication of all, or nearly all, SIVcpZP strains (e.g., Melissa in Fig. 1B and D) while 18 others were refractory to six of the eight viruses (e.g., Dona in Fig. 1 C and D). Although not all viruses could be tested in cells from all animals, 15 chimpanzees supported replication of only TAN2 and MT145 while 3 others supported replication of only EK505 and MT145 (Fig. 1D). Cells from one chimpanzee (Chip) supported the replication of MT145, EK505, GAB2, and LB715, but not MB897, BF1167, and TAN2 (Fig. 1D). These infectivity patterns were reproducible in all instances where repeat blood samples from the same individuals were available (SI Appendix, Table S1). Only a single SIVcpZP strain, MT145, was able to replicate in CD4+ T cells from all 28 chimpanzees (Fig. 1D). These results showed that, unlike human CD4+ T cells, chimpanzee CD4+ T cells are refractory to a diversity of viruses, with susceptibility influenced by both host- and virus-specific determinants.

**Results**

**Chimpanzee CD4+ T Cell Cultures Differ in Their Susceptibility to SIVcpZP Infection.** Generating infectious molecular clones (IMCs) of SIVcpZP, we previously noted that some viruses that replicated efficiently in human CD4+ T cells were unable to infect chimpanzee CD4+ T cells (44). To examine this surprising phenotype, we obtained leftover blood samples from 28 healthy chimpanzees housed at US primate centers and infected their CD4+ T cells with a panel of eight chimpanzee viruses representing both SIVcpzPr (MT145, EK505, MB897, LB715, and GAB2) and SIVcpZPs (BF1167, TAN2, and TAN13) strains (21, 44–47). SIVcpZPs were transfected, normalized based on infectivity in a permissive cell line (TZM-bi), and used to infect CD4+ T cells at a multiplicity of infection of 0.1 (Fig. 1). As observed previously (21, 44, 47), all SIVcpZP strains replicated efficiently in human CD4+ T cells (Fig. 1A). However, their ability to establish a productive infection in chimpanzee CD4+ T cells varied considerably. Of the 28 chimpanzees tested, only 9 supported the replication of all, or nearly all, SIVcpZP strains (e.g., Melissa in Fig. 1B and D) while 18 others were refractory to six of the eight viruses (e.g., Dona in Fig. 1 C and D). Although not all viruses could be tested in cells from all animals, 15 chimpanzees supported replication of only TAN2 and MT145 while 3 others supported replication of only EK505 and MT145 (Fig. 1D). Cells from one chimpanzee (Chip) supported the replication of MT145, EK505, GAB2, and LB715, but not MB897, BF1167, and TAN2 (Fig. 1D). These infectivity patterns were reproducible in all instances where repeat blood samples from the same individuals were available (SI Appendix, Table S1). Only a single SIVcpZP strain, MT145, was able to replicate in CD4+ T cells from all 28 chimpanzees (Fig. 1D). These results showed that, unlike human CD4+ T cells, chimpanzee CD4+ T cells are refractory to a diversity of viruses, with susceptibility influenced by both host- and virus-specific determinants.

**Refractory CD4+ T Cells Exhibit an Entry Block.** To determine at which step in the life cycle SIVcpZP was blocked, we generated a replication-competent SIVcpZP reporter virus by inserting an enhanced green fluorescent protein (eGFP) gene between the env and nef genes of the SIVcpZP MB897 molecular clone (SI Appendix, Fig. S2A). A transfection-derived stock of this reporter virus was then used to infect activated human and chimpanzee CD4+ T cells. Consistent with results obtained for the WT MB897 strain (Fig. 1D), eGFP expression was detected in CD4+ T cells from humans and susceptible chimpanzees, but not in CD4+ T cells from refractory chimpanzees (SI Appendix, Fig. S2B). However, complementation of the MB897 reporter virus with the vesicular stomatitis virus G protein (VSV-G) resulted in productive infection of all cultures (SI Appendix, Fig. S2B). These data indicated that SIVcpZP was inhibited at the level of cell entry.

**CD4 Polymorphisms Govern the Susceptibility of CD4+ T Cells to SIVcpZP Infection.** In addition to CD4, SIVcpZP utilizes the chemokine receptor CCR5 to infect target cells (21, 44). Since the VSV-G complementation studies suggested a block at the receptor and/or coreceptor level, we sequenced the CD4 and
CCR5 genes of all 28 chimpanzees. In contrast to the CCR5 gene, which comprises a single coding exon, the CD4 gene is expressed from nine exons, some of which have been reported to be polymorphic (42). Since previous studies did not ensure linkage of variable sites and failed to guard against PCR artifacts (42), we extracted RNA from chimpanzee CD4+ T cells and used limiting dilution RT-PCR to amplify single CD4 transcripts (48). These analyses showed that none of the 28 chimpanzees exhibited mutations in their CCR5 gene. However, analysis of their CD4 sequences revealed several single nucleotide polymorphisms (SNPs), four of which changed the amino acid sequence of the mature CD4 protein (SI Appendix, Fig. S3A).

**Fig. 1.** Chimpanzee CD4+ T cells differ in their susceptibility to SIVcpz infection. (A–C) The replication potential of eight SIVcpz strains is shown in activated CD4+ T cells from one human (A) and two chimpanzees (B and C). Viral replication was monitored in culture supernatants by determining reverse transcriptase (RT) activity (ng/mL). (D) CD4+ T cells from 28 captive chimpanzees (columns) were infected with different strains of SIVcpz (rows) and monitored for viral replication. The chimpanzee-adapted HIV-1 SG3 strain was used for positive control (SI Appendix, SI Materials and Methods). RT activity at day 13 was used to classify each culture as supporting robust (>5 ng/mL RT, dark circles), weak (1 to 5 ng/mL RT, grey circles), or no (<1 ng/mL RT, white circles) viral replication (SI Appendix, Table S1). Replication results are shown in relation to the respective chimpanzee CD4 genotype, with the position of polymorphic amino acid residues highlighted. (E) Protein sequences of five chimpanzee CD4 variants shown in comparison with human CD4 (residues are numbered according to their position in the mature CD4 protein), extracellular (D1 to D4), transmembrane (TM), and intracytoplasmic domains of the protein are indicated relative to their coding exons (highlighted by alternating black and blue text), with the D1 domain underlined in red. Dots indicate amino acid identity to the human sequence, with the polymorphic sites in the D1 domain highlighted in red. Conserved and polymorphic potential N-linked glycosylation sites (PNGSs) are shaded in gray and red, respectively.
all chimpanzee CD4 alleles (Fig. 1E). Based on the amino acids present at positions 40, 52, 55, and 68, we designated the chimpanzee CD4 alleles QNIP, QNVP, QNVT, RNVP, and RKVT, respectively (SI Appendix, Table S2); the latter two had not previously been described (42).

Comparing the CD4 allelic diversity between permissive and refractory cultures, we found a remarkable association between the CD4 genotype and SIVcpz infection phenotypes (Fig. 1D). CD4+ T cells from four chimpanzees, which were heterozygous for the QNVT and QNVP alleles, supported replication of all eight SIVcpz strains while CD4+ T cells from four other chimpanzees, which were heterozygous for QNVT and QNIP alleles, supported only two of the eight SIVcpz strains. This was also true for chimpanzees heterozygous for the QNVT and RKVT alleles except this genotype inhibited TAN2 instead of EK505. In general, CD4+ T cells from chimpanzees with the same CD4 genotype were susceptible to infection by the same set of viruses while even a single amino acid substitution in one of the two CD4 alleles changed the number or types of viruses that were able to replicate (Fig. 1D).

Chimpanzee Subspecies Differ in Their CD4 Diversity. The finding of new CD4 alleles among captive chimpanzees suggested that additional variants might exist in wild populations. To determine the full extent of CD4 diversity, we thus made use of our extensive collection of blood and fecal samples obtained previously from wild and sanctuary chimpanzees for molecular epidemiological studies of SIVcpz and ape Plasmodium infections (22, 49, 50). Samples were selected based on their geographic and subspecies origin, SIVcpz infection status, and individual information (SI Appendix, Table S3). Since the chimpanzee CD4 gene spans a 19-Kb region on chromosome 12, with a large intron (13.7 Kb) separating exons 2 and 3, we were unable to amplify the entire CD4 coding region from a single DNA template. However, since all CD4 polymorphisms were located in the D1 domain, we amplified exon 2 (247 bp) and exon 3 (222 bp) separately, sequencing their respective amplicons without fragmentation to maintain linkage between polymorphic sites. Homozygous loci were amplified up to eight times to exclude allelic dropout. Using this approach, we CD4 genotyped 60 P. t. verus, 41 P. t. ellioti, 246 P. t. troglodytes, and 197 P. t. schweinfurthii apes that were sampled at 38 sites throughout their range (SI Appendix, Fig. S1). Although exons 2 and 3 were sequenced separately, we were able to infer 13 different D1 domain haplotypes from individuals who were homozygous for one of the two exons (SI Appendix, Fig. S3B). These haplotypes contained one synonymous SNP (g/a) at the second nucleotide of exon 2 (nucleotide position 51 of the CD4 coding sequence) and five nonsynonymous SNPs, which resulted in amino acid substitutions at positions 25, 40, 52, 55, and 68 of the mature CD4 protein. Comparison of these to CD4 sequences from other ape species identified the (g)QONVP allele as the ancestral state, which appeared to have diversified both by point mutations and recombination, including within exons 2 and 3 (SI Appendix, Fig. S3B). Thus, in contrast to the human CD4, for which the most common D1 variant occurs at a frequency of only 6.6 × 10^{-5} (SI Appendix, SI Materials and Methods), the chimpanzee CD4 is highly polymorphic in this domain, with nine coding variants of CD4 identified in wild populations (Fig. 24).

Chimpanzees are believed to have originated in west central Africa where they initially split into two lineages when the ancestor of P. t. verus and P. t. ellioti diverged from the ancestor of P. t. troglodytes and P. t. schweinfurthii (51). Given these relationships and the fact that only central and eastern chimpanzees are SIVcpz-infected, we expected P. t. verus and P. t. ellioti to exhibit similar levels of CD4 diversity and to differ from P. t. troglodytes and P. t. schweinfurthii both in the number and distribution of CD4 alleles. However, this was not the case. Instead, we found P. t. ellioti and P. t. troglodytes to be most similar, with exon 2 QQ and QR alleles and exon 3 NVP, NIP, and KVT alleles present at comparable frequencies in both subspecies. Although P. t. troglodytes exhibited two additional exon 3 alleles (NVT and KVP), both were extremely rare and may thus have been missed in the less extensively sampled P. t. ellioti subspecies. P. t. schweinfurthii also encoded multiple CD4 alleles, but their frequencies differed markedly from those in P. t. troglodytes and P. t. ellioti. For example, using the CD4 exon 3 haplotype frequencies, G_ST (a measure of population differentiation) (52) between P. t. troglodytes and P. t. schweinfurthii was 0.31, compared with only 0.11 between P. t. troglodytes and P. t. ellioti. Moreover, none of the eastern chimpanzees encoded the R40 allele, but instead some encoded an R25 allele, which was found only in chimpanzees living in Gombe National Park in Tanzania (SI Appendix, Table S3). The most surprising finding was that P. t. verus apes did not exhibit any CD4 polymorphisms. Although 60 individuals were sampled at five different locations throughout West Africa (SI Appendix, Fig. S1), every single one was homozygous for the (a)QONVT allele (Fig. 2B). Given the scarcity of the exon 3 NVT allele in the other three subspecies, the predominance of this allele is unlikely the result of a founder effect. Instead, P. t. verus apes appear to have undergone a selective sweep, possibly in response to an SIV-like pathogen that has since gone extinct. Taken together, these findings suggest that the various CD4 polymorphisms evolved before the introduction of present-day SIVcpz and likely already existed in the chimpanzee ancestor.

CD4 Polymorphisms Inhibit SIVcpz Infection. To examine the interaction of the nine CD4 alleles with individual SIVcpz Env variants, we developed a single round infection assay by transfecting 295T cells with vectors expressing human and chimpanzee CD4 alleles, along with the chimpanzee CCR5 gene, and then infecting these cells with pseudotyped viruses carrying the Env of different SIVcpz strains. As the pseudotyping backbone, we used an env-deficient version of the GFP-expressing MB897 reporter virus, which allowed the use of flow cytometry to identify infected (GFP-expressing) cells.
Chimpanzee CD4 Polymorphisms Interfere with Env Binding by Altering Contact Residues at the CD4–Env Interface. Having identified protective CD4 alleles, we next sought to examine the breadth of this protection and the associated mechanism(s). Both R25- and R40-containing CD4 alleles inhibited primarily SIVcpzPs strains (Figs. 3 and 4), but CD4 genotyping of 128 chimpanzees from Gombe National Park identified two SIVcpzPts-infected individuals that were R25-heterozygous. Since the virus of one of these individuals (TANI) has been molecularly cloned (44), we tested its Env in the CD4 transfection assay (Fig. 4A). This analysis showed that TANI, in contrast to the other SIVcpzPts Env, was able to utilize the R25 allele (Fig. 4A). Similarly, of five SIVcpzPt strains, four were resistant to the R40 inhibition, and all four were derived from chimpanzees that encoded at least one R40-containing allele (Fig. 4C). Thus, for both R25 and R40 polymorphisms, in vivo adaptation had generated SIVcpz strains that were capable of utilizing these CD4 variants.

To identify the mechanisms of inhibition mediated by R25 and R40, we mapped these CD4 residues onto an existing crystal structure of the human CD4 bound to HIV-1 gp120 (38). We then searched for Env amino acids that were in close proximity to these CD4 residues but differed between restricted and permissive SIVcpz strains (Figs. 4B and D). For the R25 polymorphism, we found that Env residue 474 (HXB2 reference sequence numbering) was negatively charged or neutral in all viruses that were able to utilize the RQNVP allele, but positively charged in SIVcpz strains that were unable to utilize this allele (Fig. 4B). This finding suggested that the CD4 R25 repulsed positively charged Env amino acids at position 474, thus destabilizing the Env–CD4 interaction. Since the permissive TAN1 strain encoded a valine at this position, we mutagenized a restricted Env (TAN2) by replacing its lysine with a valine (K474V). This single amino acid substitution was sufficient to render the TAN2 Env infectious for cells expressing the RQNVP allele without altering its ability to utilize the QONOVP allele (Fig. 4E). Similarly, residue 455 tended to be negatively charged in Env that could use R40-containing CD4 alleles, but neutral in Env that were blocked by this allele (Fig. 4D).

In this instance, the different orientation of the R40 side chain was incompatible with an interaction with Env residue 283, the contact residue for CD4 Q40. However, this was postulated to be the tip of a long-range salt bridge between CD4 R40 and a negatively charged residue at Env position 455. In line with this prediction, replacing a threonine at position 455 with an aspartic acid (T455D) in the restricted MB897 Env allowed SIVcpz infection without altering its ability to utilize the QONOVP allele (Fig. 4E).

To examine the effect of the T455D mutation in the context of a replication-competent virus, we introduced this change into the MB897 infectious molecular clone and tested its infectivity in CD4+ cells from a chimpanzee (Chip). This was heterozygous for the QONOVP and QONOVT alleles (SI Appendix, Fig. S7A). Although Env amino acid 455 is unlikely the only residue impacted by the R40 substitution (the LB715 Env encodes a T at position 455 but is able to use the QONOVP allele), it seems clear that changes in the charge of Env/CD4 contact residues represent one mechanism by which CD4 polymorphisms prevent Env-mediated cell entry. Of note, R25 and R40 blocked CD4/Env interaction by different means since Env mutations that restored infectivity for one of these CD4 alleles did not restore infectivity for the other (Fig. 4E).

Chimpanzee CD4 Polymorphisms Cause Glycan–Glycan Steric Hindrance. In contrast to the R25 and R40 alleles, the T68 polymorphism inhibited all SIVcpz Env regardless of their baseline infectivity (Fig. 3). Since T68 creates a PNGS at position N66 in addition to the N32 that is present in all chimpanzee CD4 alleles, we asked whether these sites were indeed glycosylated. Immunopurification of CD4 proteins from cells transfected with human and chimpanzee (QONOVP and QONOVT) alleles revealed differences in electrophoretic mobility consistent with glycan occupancy, which was confirmed by endoglycosidase H treatment (SI Appendix, Fig. S8). To determine whether these D1 domain glycans interfered
glycan clashes were env genes from several SIV lineages into expression T cells from QQNVT- N66 T cells from one, but not and N66 Steric hindrance between CD4- and SIVcpz Env-encoded glycans. (E) T cell from a homozygous chimpanzee encoding QQNVP and MB897 N460Q *F and | and A and position 40 in C circles represent average infectivity 0.01, unpaired t test). (43x66) both QQNVP- and QQNVT-expressing cells (Fig. 5 and the triple mutant increased the infectivity of the MB897 Env in transient CD4 transfection and replication studies, albeit not to the level of other QQNVT-permissive SIVcpz strains, such as MT145. Removal of the other two glycans failed to restore infectivity, although this was not due to a fitness cost, since all mutants replicated to the same extent in human CD4 T cells (SI Appendix, Fig. S7B). Thus, in silico modeling of the CD4 and Env glycan shields did not fully recapitulate the extent of their interactions.

Chimpanzee CD4 Polymorphisms Protect Against SIVs Infecting Old World Monkeys. To test whether the inhibitory effect of the chimpanzee CD4 polymorphisms extended to more diverse SIVs, we cloned env genes from several SIV lineages into expression vectors, generated pseudoviruses, and tested their functional integrity in TZM-bl cells. Infectious Envs were then tested for CD4 dependence, and only those that required CD4 for cell entry with glycans on the various SIVcpz Envs, we modeled their position in the crystal structure of the human CD4 bound to HIV-1 gp120 (38). This analysis identified three glycans in the MB897 Env that were predicted to clash with glycans in the chimpanzee CD4 protein (Fig. S4). Env glycan N460 was predicted to clash with CD4 glycan N32 while Env glycans N295 and N446 were predicted to clash with glycans in the chimpanzee CD4 (38). Polymeric CD4 residues are shown in dark blue (Q) and red (R), respectively.

A model of two SIVcpz Envs (MB897, TAN2) in complex with the chimpanzee CD4 QQNVT variant (containing both N32 and N66 glycans) is shown. Env residues are shown in purple, Env glycans in green, and CD4 glycans in red. (D) Infectivity of WT and mutant MB897 Envs lacking the indicated glycans in cells expressing different CD4 alleles (parentheses indicate the number of D1 domain glycans). Asterisks denote significant enhancement of the mutant Env relative to the MB897 WT (orange) for that particular allele. (E) Replication of wild-type and mutant MB897 viruses in CD4+ T cells from one heterozygous chimpanzee encoding QQNVP and QQNVT alleles (C) and two homozygous chimpanzees encoding the QQNVT allele (D and E).
were further analyzed. Using the transient CD4 transfection assay, we identified nine Envs from SIVs infecting mustached monkeys (Cercopithecus cephus), H'ooest's monkeys (Cercopithecus hooesti), western red colobus (Procolobus badius), red-tailed monkeys (Cercopithecus ascanius), sooty mangabeys (Cercocebus aethiops), and African green monkeys (Chlorocebus talapoin) that mediated entry into cells expressing the human CD4 (Fig. 6A). However, when these same Envs were tested in cells expressing the chimpanzee QONVP allele, the great majority exhibited markedly reduced infectivity, with some of them inhibited even more in cells expressing the chimpanzee QONVT allele (SI Appendix, Fig. S6F). To examine possible mechanisms, we mutagenized the conserved N32 glycan in both the QONVP and QONVT alleles. Changing the asparagine at position 32 to a glutamine (N32Q) in the QONVP allele enhanced the infectivity of most SIV Envs, suggesting that their inhibition was primarily glycan-mediated (SI Appendix, Fig. S6F). However, the same mutation in the QONVT allele did not restore Env infectivity, suggesting that one glycan in the D1 domain is sufficient to exert inhibition (Fig. 6B and SI Appendix, Fig. S6F).

For six SIV Envs that retained infectivity for the chimpanzee QONVP allele, we also tested the remaining CD4 polymorphisms (Fig. 6C). Both the N52K and V55I polymorphisms inhibited Envs of the SIVsmm lineage but had little effect on the other SIV Envs, except for SIVagm, which was enhanced by the K25 substitution (SI Appendix, Fig. S6G). Since the African green monkey CD4 encodes an R at position 52, this substitution likely generated a chimpanzee CD4 allele that resembled the cognate SIVagm receptor more closely. The R25 substitution inhibited one SIVagm and one SIVsmm Env, but none of the other SIV Envs (SI Appendix, Fig. S6G). The most pronounced protective effect was again observed for the R40 substitution, which reduced cell entry of all SIV Envs by ~10-fold, except for the Env of SIVasc (SI Appendix, Fig. S6G). Thus, like for SIVcpz strains, CD4 polymorphism-mediated inhibition of SIV cell entry was strain-specific and context-dependent.

CD4 Receptor Diversity Protects Wild Chimpanzees from SIVcpz Infection. Finally, we asked whether the various chimpanzee CD4 alleles influenced SIVcpz acquisition in vivo. To address this, we selected chimpanzee communities at two field sites where CD4 diversity could be analyzed in the context of high SIVcpz prevalence rates (SI Appendix, Table S3). One included Gombe National Park, where members of the Mitumba, Kasekela, and Kalandé communities have been monitored for SIVcpz for two decades and where the negative impact of this infection on chimpanzee health was first demonstrated (27, 28). The other included the Lobéké (LB) and Mambélé (MB) area in southeastern Cameroon where chimpanzees harbor the closest relatives of HIV-1 group M at high prevalence rates (22). Using logistic regression to examine the effect of CD4 substitutions on SIVcpz infection relative to the ancestral CD4 allele, we found a suggestion of a protective effect for the R25 polymorphism in the Gombe population although this did not reach statistical significance (Fig. 7A). The 70 chimpanzees from the Lobéké (LB) and Mambélé (MB) area lacked the R25 polymorphism but encoded CD4 alleles with other substitutions in exon 2 and exon 3. When these were analyzed, none of them reached statistical significance although there was a trend for R40 and K52 to be associated with lower, and for I55 to be associated with higher, SIVcpz infection rates (Fig. 7B). Surprisingly, the T68 substitution did not appear to have a detectable effect. However, since only 4 of 46 individuals encoded T68 in the absence of K52, the impact of these two substitutions could not be differentiated. Interestingly, when the LB/MB data were analyzed at the two exon level, the exon 3 allele was associated with 3.4-fold lower odds of being infected ($P = 0.03$, 95% CI: 1.07-10.8-fold) (Fig. 7). These data thus suggest that K52 (alone or in combination with T68) and, possibly, also R25 protect wild chimpanzees against infection by locally circulating SIVcpz strains although definitive conclusions must await more extensive sampling of wild populations.

Discussion

Pathogenic SIVs have long been assumed to be the driving force behind the positive selection of primate CD4, but direct evidence has been lacking (39, 40). Here, we show that wild chimpanzees encode nine different CD4 variants, all of which are the result of nonsynonymous SNPs in the D1 domain that binds the HIV/SIV Env trimer (Figs. 1 and 2). Testing their impact on virus infection, we found that these polymorphisms inhibited cell entry of SIVcpz, as well as more distantly related SIVs, albeit in a strain-specific and context-dependent manner (Figs. 3 and 7). The most striking effects were seen for amino acid substitutions at positions 25, 40, and 68 of the mature CD4 protein in codons not previously identified to be under positive selection (39, 40). In silico modeling and site-directed mutagenesis identified charge changes at the Env/CD4 interface (Fig. 4) as well as clashes between Env- and CD4-encoded glycans (Fig. 6) as mechanisms of cell entry inhibition. The finding of polymorphisms in the chimpanzee CD4 that alter its function as a virus receptor, but do not appear to interfere with its MHC class II binding ability, strongly suggest that they evolved in response to SIV-mediated selection.

Although SIVcpz is pathogenic in chimpanzees (26–28), it is unlikely that it is the only SIV that has exerted pressure on the chimpanzee CD4 protein. *P. t. verus* and *P. t. ellioti* do not harbor SIVcpz, which suggested a more recent introduction of this virus, after the split of these subspecies from *P. t. schweinfurthii*. Indeed, the structure of the SIVcpz genome represents a complex mosaic, which resulted from the cross-species transmission and recombination of SIV lineages infecting red-capped mangabeys (Cercocebus torquatus) and certain Cercopithecus monkey species (10). Since the ranges of these species overlap that of *P. t. troglodytes* apes, it seemed likely that SIVcpz
Effect of CD4 polymorphisms on SIVcpz infection rates in wild chimpanzees. A logistic regression was used to estimate the effects of CD4 polymorphisms on SIVcpz infection in wild-living chimpanzees from (A) Gombe and (B) Lobéké/Mambébé. Dots indicate the estimated effect of the presence of the given residue (Left) or allele (Right) relative to the ancestral state, with horizontal lines indicating the 95% confidence intervals (substitutions/alleles that could not be estimated by the model due to insufficient variation in the population are not shown). Dashed vertical lines mark a fold change of 1, indicating no predicted change in SIVcpz infection rate relative to the ancestral state.

first emerged in west central Africa and subsequently spread eastward (16). However, the fact that P. t. ellioti, P. t. troglodytes, and P. t. schweinfurthii share several CD4 alleles indicates that CD4 diversification preceded their divergence (Fig. 2). Thus, either current-day SIVcpz is much older than previously thought and selected CD4 variants in ancestral chimpanzees before becoming extinct in P. t. verus and P. t. ellioti, or chimpanzees have been episodically infected with different SIVs throughout their evolutionary history, which placed pressure on their CD4. Still another possibility is that some, or all, of the CD4 polymorphisms were selected by unrelated pathogen(s) (53, 54). However, others have independently implicated an ancient SIV infection to explain the skewed MHC class I repertoires of chimpanzees and bonobos, which include allotypes that resemble AIDS-protective human HLA alleles (55). Thus, there are several lines of evidence to suggest that different SIVs, both extant and extinct, have shaped the evolution of the chimpanzee CD4.

The finding of only a single CD4 allele in western chimpanzees is intriguing. This could indicate that the selection pressure that maintains CD4 diversity in the other subspecies has been removed although these apes are almost certainly exposed to a similar variety of SIVs from various sympatric monkey species. Alternatively, exposure to new SIV(s) as the ancestors of other three subspecies, but its fixation in the western subspecies, which has been linked to its catastrophic population decline (28), may be the absence of the protective RQNVP allele. Although it is unclear why the distribution of the RQNVP allele is so different between the otherwise interconnected Gombe communities, it will be important to determine whether its frequency can explain differences in SIVcpz prevalence rates that are frequently pronounced among P. t. schweinfurthii communities in the easternmost part of their range (21).

Humans acquired the ape precursors of HIV-1 by cross-species transmission on four independent occasions, resulting in groups M, N, O, and P (16). In contrast, SIVcpz appears to have arisen only once, despite frequent exposure of chimpanzees to SIVs (31). Here, we show that receptor glycosylation serves as a potent barrier to SIV cell entry. Both conserved (N32) and variable (N66) D1 domain glycans were shown to inhibit not only SIVcpz, but also other SIV strains (Fig. 6). While many additional SIV lineages remain to be analyzed, our data indicate that steric hindrance between cell entry receptor-encoded and virus surface protein-encoded glycans represents a mechanism of antiviral protection that has not previously been described. Indeed, glycosylation of the CD4 receptor is common among primates where PNGS are found at variable positions in the D1 domain (SI Appendix, Table S5). Moreover, all of these appear to be under positive selection within the primate lineage (SI Appendix, Fig. S9). Humans lack D1 domain glycans, which may explain their relative susceptibility to SIVcpz, SIVgor, and SIVsmm infections, which have crossed the species barrier on at least 12 occasions (16). Conversely, the presence of glycans on the chimpanzee CD4 must be one reason why chimpanzees are largely resistant to experimental HIV-1 infection. It will be interesting to examine the evolution of the Env glycan shield in HIV-1 strains that were able to establish a productive infection in chimpanzees after multiple rounds of in vivo adaptation (65).

The discovery of CD4-mediated protection has practical implications for AIDS vaccine development. We recently discovered that a subset of SIVcpz strains share unexpected antigenic cross-reactivity with HIV-1 in the functionally important V1V2 region of the Env trimer apex (45). This finding raised the question whether SIVcpz Env, which are otherwise antigenically highly divergent from HIV-1, could serve to immunofocus B cell responses in humans to this critical epitope. Indeed, a minimally
modified Env of the SIVcpz strain MT145 was recently shown to display selective binding to HIV-1 V2-apex broadly neutralizing antibodies (bNabs) and their precursors, and to prime heterologous (tier 2) neutralizing antibody responses in V2 apex bNab precursor antibody-expressing knock-in mice (66). MT145 is the only SIVcpz strain that was able to replicate in CD4+ T cells of all chimpanzee lines (Fig. 1), at least in part because it lacks Env glycans predicted to clash with N32 and N66 in the chimpanzee CD4 (Fig. S4). Indeed, cryo-EM analysis of the MT145 Env trimmer revealed that its structure is remarkably similar to that of the HIV-1 Env, except for a shift in the arrangement of its glycans (66). It thus appears that MT145 has evolved to accommodate the chimpanzee CD4-mediated cell entry block by rearranging its Env glycan shield. However, the absence of the N460 glycan exposes a long V5 loop (Fig. S4), which has the potential to induce unwanted (off-target) antibody responses. Since there are other SIVcpz Env that contain the cross-reactive V2 apex epitope but lack the long unshielded V5 loop (45), they may be more suitable components of an immunofocusing strategy to elicit V2 apex bNab responses. As vaccines including SIVcpz Env immunogens are moving toward human clinical testing, understanding the impact of chimpanzee CD4 diversification on their structure and function will inform AIDS immunogen design.

Methods

Vectors. The construction and biological characterization of the SIVcpz IMG have previously been reported (21, 44–47). Insertion of a GFP-internal ribosome entry site (IRES) cassette between the env and nef genes of the MB897 IMC generated a replication-competent SIVcpz-GFP reporter virus, which was further modified by inserting a frameshift at position 6,493 in the env gene for pseudotyping studies. WT or codon-optimized SIVcpz and SIV Env genes were cloned into pcDNA3.1. Full-length chimpanzee and human CD4 cDNA coding sequences, as well as the chimpanzee CCR5 gene, were cloned into pMSCVpuro (Takara Bio Inc.).

Virus Stocks. Viral stocks were generated by transfecting of 293T cells with SIVcpz IMGc and by testing the culture supernatants for infectivity on TZM-bl cells. Pseudovirus stocks were generated by cotransfecting the MB897ΔEnv-GFP backbone with WT or codon-optimized SIVcpz and SIV Env expression plasmids and also titered on TZM-bl cells. The replication-competent MB897ΔEnv-GFP IMC was used for VSV-G complementation studies.

CD4+ T Cell Cultures. Blood samples were obtained from captive chimpanzees housed at the Yerkes National Primate Research Center, the Southwest National Primate Research Center, and the New Iberia Research Center during their annual health examination. Only leftover material was used, which was approved by the respective Institutional Animal Care and Use Committees. Human blood was purchased (ZenBio, Inc). Chimpanzee and human CD4+ T cells were isolated, activated, and infected overnight at a multiplicity of infection (MOI) of 0.1 as described (44, 67). Virus replication was assessed by monitoring reverse transcriptase activity (Sigma-Aldrich) or the presence of p24 core protein (AlphalISA Detection Kit; Perkin-Elmer) in culture supernatants.

CD4 and CCR5 Genotyping. To determine the CD4 and CCR5 genotype of captive chimpanzees, total RNA was extracted from activated CD4+ T cells and reverse transcribed using SuperScript III (Thermo Fisher) gene-specific primers. To preclude PCR artifacts, cDNA was endpoint diluted (48) to amplify single mRNA templates (see SI Appendix, SI Materials and Methods for primer sequences and amplification conditions). Multiple amplicons were MiSeq sequenced to determine the CCR5 and CD4 genotype (SI Appendix, Table S2). To determine chimpanzee wild chimpanzee populations, samples were selected from existing specimen banks (4, 21, 22, 27, 28, 49, 50) based on geographic origin, subspecies association, SIVcpz infection status, and host mitochondrial and microsatellite information, as well as sample availability and quality (SI Appendix, Tables S2 and S5 for available sample information). CD4 exon 2 (247 bp) and exon 3 (222 bp) regions were amplified using primers in adjacent introns (SI Appendix, SI Materials and Methods). Amplicons were MiSeq sequenced without fragmentation to ensure linkage of variable sites (68, 69). Homozygous loci were amplified at least eight times to exclude allelic dropout (SI Appendix, Table S3).

**Transient CD4 Transfection Assay.** To determine the ability of SIVcpz and SIV Env to utilize different CD4 alleles for cell entry, we developed a single round infectivity assay. Briefly, 293T cells were cotransfected with human and chimpanzee CD4 and chimpanzee CCR5 expression plasmids, cultured for 48 h, and then plated in 96-well plates at a density of 2 × 10^4 cells per well. After 24 h, transfected cells were infected with 5,000 infectious units (IU) of SIV Env bearing pseudovirus by spinoculation, cultured for 48 h, and then analyzed for GFP expression by flow cytometry (SI Appendix, Fig. S4). CD4 and CCR5 expression levels were determined at the time of infection using relevant species-specific antibodies, and the percentage of infected cells was calculated by dividing the number of GFP-expressing cells by the number of CD4- and CCR5-expressing cells for each transfected cell population. All Env were analyzed in triplicate on three independent occasions.

**Modeling Glycosylated SIVcpz Env Trimer and CD4.** To obtain models of SIVcpz trimers and CD4 variants, the structure of HIV-1 clade G x1193.c1 SOSIP.665 (PDB ID code 5FYD) protein (70) and the human CD4 (PDB ID code 1GCI) protein (38) were used as templates in SwissModel (71). N-linked Man-5 glycans were added to both SIVcpz trimmer and CD4 models in silico using the in-house program Glycysulator at all possible glycosylation sequons. Complexes of modeled glycosylated SIVcpz and SIVcpz trimers were built by structural alignment of gp120 from 1GCI to the corresponding residues in the trimeric SIVcpz gp120. Gyan-ryan sugar spatial loops were evaluated by visual inspection of multiple glycans rotamers.

**Statistical Analyses.** To assess the effects of CD4 polymorphisms on SIVcpz- and SIV Env-mediated cell entry in vitro, we used a hierarchical Bayesian model (SI Appendix). For each pair of CD4 alleles, the fold change in infectivity for the various Env, as well as an overall effect estimate, was modeled using Stan (72). To assess the effects of CD4 polymorphisms on SIVcpz infection in vivo, we selected two different populations of eastern (Gombe) and central ( Lobèké/Mambélé area) chimpanzees that were infected at high SIVcpz prevalence rates. We then performed a logistic regression of SIVcpz status on sets of indicator variables that specified whether a chimpanzee possessed a given CD4 substitution or allelic variant, while setting the ancestral state as the background. This analysis tested whether the presence of a given substitution or allele within a community positively or negatively associated with SIVcpz infection relative to the ancestral (g) QQNV genotype although the model did not account for zygosity or synergistic effects between polymorphisms.

**ACKNOWLEDGMENTS.** We thank the staff of the Yerkes and Southwest National Primate Research Centers, as well as Jane Fontenot, Melaney Musso, and Francois Villinger (New Iberia Research Center) for providing leftover blood samples from captive chimpanzees; the staff of the Sanaga-Yong Rescue Center, the Chimpanzee Health and Rehabilitation Centre (CHRC) in Cameroon, and the Chimpanzee Rehabilitation Centre for preserving leftover blood samples from rescued apes; the staff of project PRESCIA, the World Wildlife Fund for Nature (WWF)/Democratic Republic of the Congo (DRC), and researchers from the Parc National de la Lope for fieldwork in Cameroon; the DRC, and Gabon, respectively; the Jane Goodall Institute field staff at the Gombe Stream Research Centre, as well as field assistants from the Greater Mahale Ecosystem Research and Conservation Project (GMER) for collecting chimpanzee fecal samples in Tanzania; the Gashaka Primate Project for collecting chimpanzee fecal samples in Nigeria; the Ministries of Environment and Forests and Research, Office des Parcs et Réserves, Tai National Park, and the Swiss Centre for Scientific Research for permission to perform studies in Côte d’Ivoire; the Cameroonian Ministries of Health, Forestry and Wildlife, and Scientific Research and Innovation for permission to collect samples in Cameroon; the governments of Guinea, Guinea-Bissau, Nigeria, and Gabon for permission to collect fecal samples from wild chimpanzees; the Water and Forest Ministry for permission to collect samples in the Central African Republic; the Ministry of Forest Economy and Sustainable Development for permission to collect samples in the Republic of Congo; the Ministry of Scientific Research and Technology, the Department of Ecology and Management of Plant and Animal Resources of the University of Kisangani, the Ministries of Health and Research and Innovation Ethics Committee for permission to collect samples in the DRC; and the Tanzania Commission for Science and Technology (COSTECH), the Tanzania Wildlife Research Institute (TAWIRI), and the Tanzania National Parks Authority (TANAPA) for their support and permission to conduct research in Gombe and the Greater Mahale Ecosystem (GME). This work was supported in part by National Institutes of Health Grants R01 AI120810, R01 AI50529, R37 AI050529, UM1 AI010645, UM1 AI0126620, and P30 AI045008 (to B.H.J.); Yerkes National Primate Research Center Grant P51 RR000165; Swiss National Science Foundation Grant P51 RR013986 and P51 OD011133; Agence Nationale de Recherche pour le SIDA Grant ANRS 12325 (to M.P.); and Canadian Institutes of Health Research (CIHR) Foundation Grant 352417 and Canada Research Chair CRC RCH506235 (to A.F.). R.M.R. is supported by Training Grant T32 AI 007632. S.D. is supported by a Fonds de Recherche Santé Québec (FRSQ) postdoctoral fellowship.
1. Peeters, M., Ma, D., Liegeois F., Apetrei C. (2014) Simian immunodeficiency virus infec-


4. Aghokeng AF, et al. (2010) Extensive survey on the prevalence and genetic diversity of

5. SIVs in primates in bushmeat provides insights into risks for potential new cross-species


8. with a vpu gene from greater spotted monkeys (Cercopithecus nictitans) provides new


10. Santiago ML, et al. (2003) Foci of endemic simian immunodeficiency virus infection in


14. Compton AA, Malik HS, Emeran M (2013) Host gene evolution traces the evolu-


17. chimpanzees against lentiviruses from monkeys. PloS Pathog 11:e1005149.


19. analysis of wild primate TRIMs reveals the ancient emergence of primate lentiviruses


21. Aghokeng AF, et al. (2007) Full-length sequence analysis of SIVmus in wild pop-

22. ulations of mustached monkeys (Cercopithecus cephus) from Cameroon provides evidence


26. monkey CCR5 cluster at critical sites for infections by simian immunodeficiency virus


36. monkey CCR5 cluster at critical sites for infections by simian immunodeficiency virus


41. 329:1487–1488.


46. monkey CCR5 cluster at critical sites for infections by simian immunodeficiency virus


51. 329:1487–1488.


56. monkey CCR5 cluster at critical sites for infections by simian immunodeficiency virus


61. 329:1487–1488.


66. monkey CCR5 cluster at critical sites for infections by simian immunodeficiency virus


71. 329:1487–1488.


76. monkey CCR5 cluster at critical sites for infections by simian immunodeficiency virus

