
CD4 receptor diversity in chimpanzees protects against SIV infection

http://researchonline.ljmu.ac.uk/10006/

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)


LJMU has developed LJMU Research Online for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

http://researchonline.ljmu.ac.uk/
For more information please contact researchonline@ljmu.ac.uk

http://researchonline.ljmu.ac.uk/
CD4 receptor diversity in chimpanzees protects against SIV infection

Frederic Bibollet-Ruche1,a, Ronnie M. Russell1,b,1, Weimin Liu1, Guillaume B. E. Stewart-Jones1, Scott Sherrill-Mix2,b, Yingli Li1, Gerald H. Learn1, Andrew G. Smith1, Marcos V. P. Gondim1, Lindsey J. Plenderleith1,e,e Julie M. Decker1, Julian E. Eslick3, Katherine S. Wetzel4, Ronald G. Collman5,b, Shilei Ding5,b, Andres Finzi5,b, Ahdijo Ayoubal, Martine Peeters6, Fabian H. Leendertz1, Joost van Schijndel7, Christian Goedmakers7, Els Ton1, Christophe Boesch1, Hjalmar Kuehl1, Mimi Arandjelovic1, Paula Dieuguez1, Mizuki Mura1, Chistelle Colin8, Kathelijne Koops9, Sheri Spede9, Mary K. Gonder9, Martin N. Muller1, Crickette M. Sanz2,a,b and Beatrice H. Hahn1,a,b,2

1Department of Medicine, University of Pennsylvania, Philadelphia, PA 19104; 2Department of Microbiology, University of Pennsylvania, Philadelphia, PA 19104; 3Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892; 4Institute of Evolutionary Biology, University of Edinburgh, EH9 3FL Edinburgh, United Kingdom; 5Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294; 6Department of Medicine, University of Pennsylvania at Birmingham, Birmingham, AL 35294; 7Département de Microbiologie, Infectiologie et Immunologie, Centre de Recherche du Centre Hospitalier de L’Université de Montréal, Montréal, QC H2X 0A5, Canada; 8Department of Microbiology, Infectiologie et Immunologie, Université de Montréal, Montréal, QC H2X 0A5, Canada; 9Département Translationnelle Appliquée au VIH et aux Maladies Infectieuses, Institut de Recherche pour le Développement, University of Montpellier, INSERM, 34090 Montpellier, France; 10Research Group Epidemiology of Highly Pathogenic Microorganisms, Robert Koch Institut, 13353 Berlin, Germany; 11Department of Primatology, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany; 12Chimbo Foundation, 1011 PW Amsterdam, the Netherlands; 13Projet Primates France, Centre de Conservation pour Chimpanzés, BP 66, 30016 Avignon, France; 14Department of Anthropology, University of Zurich, CH-8006 Zurich, Switzerland; 15Sanaga-Yong Chimpanzee Rescue Center, Defense of Animals-Africa, Portland, OR 97204; 16Department of Biology, Drexel University, Philadelphia, PA 19104; 17Department of Anthropology, University of New Mexico, Albuquerque, NM 87131; 18Department of Anthropology, Washington University in St. Louis, St. Louis, MO 63130; 19Kongo Program, Wildlife Conservation Society, BP 14537 Brazzaville, Republic of the Congo; 20Lester E. Fisher Center for the Study and Conservation of Apes, Lincoln Park Zoo, Chicago, IL 60614; 21Chimpanzee Research Center, The Jane Goodall Institute-Congo, BP 1206 Pointe Noire, Republic of Congo; 22Africa Programs, The Jane Goodall Institute, Vienna, VA 22182; 23School of Natural Sciences and Psychology, Liverpool John Moores University, L3 3AF Liverpool, United Kingdom; 24Department of Ecology and Management of Plant and Animal Resources, Faculty of Sciences, University of Kisangani, BP 2012 Kisangani, Democratic Republic of the Congo; 25Gombe Stream Research Center, The Jane Goodall Institute, Kigoma, Tanzania; 26Department of Psychology, Franklin and Marshall College, Lancaster, PA 17604; and 27Department 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 repulsion 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 genotypes 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 the canonical N-linked glycosylation site. One allele was associated with lower SIVcpz prevalence rates in the wild. These results indicate that substitutions in the D1 domain of the chimpanzee CD4 can prevent SIV cell entry. 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.

Human and simian immunodeficiency viruses (HIV/SIVs) represent a diverse group of lentiviruses that infect over 40 primate species in sub-Saharan Africa (1). Most of these comprise Old World monkeys (Cercopithecidae), but chimpanzees (Pan troglodytes) and western gorillas (Gorilla gorilla) also harbor SIV (2–5). Studies of antiviral restriction factors, such as APOBEC, tetherin, and others SIVs to which they are exposed.

Significance

CD4 is known to have evolved rapidly in primates, but the reason for this diversification is unknown. Here, we show that polymorphisms in the simian immunodeficiency virus (SIV) envelope (Env) binding domain of the CD4 receptor modulate the susceptibility of chimpanzee CD4+ T cells to SIV infection by interfering with Env–CD4 interactions required for viral entry. Both amino acid substitutions and N-linked glycosylation sites in the D1 domain blocked Env-mediated entry of a number of SIVs, including viruses that infect primates on which chimpanzee prey. These data identify steric hindrance between cell entry receptor-encoded and virus surface protein-encoded glycans as a mechanism of antiviral protection and suggest that selection pressures by primate lentiviruses, both extant and extinct, have shaped the evolution of chimpanzee CD4.

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 repulsion 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 genotypes 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 the canonical N-linked glycosylation site. One allele was associated with lower SIVcpz prevalence rates in the wild. These results indicate that substitutions in the D1 domain of the chimpanzee CD4 can prevent SIV cell entry. 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.
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 primates 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 (Chlorocebus genus) and sooty mangabeys (Cercocebus atys), 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 ellioti), 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 genome mosaic showed that SIVcpz acquisition by cross-species transmission and recombination of SIVs infecting monkey species on which chimpanzees prey (10). Central chimpanzees subsequently transmitted SIVcpzPtt to humans, generating both pandemic (group M) and nonpandemic (group N) HIV-1, as well as to western lowland gorillas (Gorilla gorilla gorilla), generating SIVgor (5, 22, 23). Western gorillas, in turn, transmitted SIVgor to humans, generating HIV-1 groups O and P (24, 25). Although the impact of SIVgor infection on the health and longevity of wild gorillas is not known, SIVcpz is pathogenic in its natural host (26, 27). Infected chimpanzees in Gombe National Park, Tanzania, have a higher mortality than uninfected chimpanzees (27, 28). In addition, infected females have lower birth rates and a higher infant mortality (27, 28). Even members of chimpanzee subspecies that do not naturally harbor SIVcpz are susceptible to infection and disease. Examples include the transmission of SIVcpzPtt from a central chimpanzee to a Nigeria-Cameroonian cage mate (29) and the experimental infection of a western chimpanzee with SIVcpzPts that resulted in high titer viremia, CD4 T cell depletion, and clinical AIDS requiring antiretroviral therapy (30).

Despite their genetic diversity, all SIVcpz strains characterized to date share an identical genome structure (10). This indicates that SIVcpz arose only once, which may seem surprising, given that chimpanzees are routinely exposed to a plethora of SIVs through their hunting behavior (31). The absence of additional SIV infections has been attributed to the antiviral activity of innate restriction and viral dependency factors, such as proteins of the APOBEC3 family and the nucleoporin RanBP2, which represent potent barriers to cross-species transmission (7, 32). However, these host factors cannot explain the uneven distribution of SIVcpz in wild chimpanzee populations, which is characterized by high prevalence rates in some communities and rare or absent infection in others (4, 21, 22). It is also unclear why Nigeria-Cameroonian chimpanzees, which are susceptible to infection (29), do not harbor SIVcpz. Although separated from SIVcpz-infected P. t. troglodytes apes by the Sanaga River, this boundary is not absolute (22, 33). Chimpanzees thus appear to have evolved additional protective mechanisms that limit their infection with SIVcpz and other SIVs.

Unlike lentiviruses infecting other mammals, SIVs gain entry into target cells by using CD4, which is expressed on a variety of immune cells, including helper T cells, macrophages, and dendritic cells. Helper T cells require CD4 to stimulate the interaction of their T cell receptor (TCR) with major histocompatibility complex class II (MHC II) molecules expressed on antigen-presenting cells. As part of the TCR complex, the most outward domain of CD4 (D1 domain) interacts with a nonpolymorphic region on MHC II (34–36). Interestingly, this same D1 domain is also the region that is bound by the envelope (Env) glycoprotein of primate lentiviruses (37, 38). Several groups have compared the amino acid sequences of CD4 between different primate species and found that residues in the D1 domain are under positive selection (39, 40). Moreover, African green monkey CD4. T cells, macrophages, and chimpanzees are known to encode polymorphic CD4 receptors (41–43). It has thus been suggested that the CD4 diversification in the primate lineage is the result of SIV-driven selection (40); however, evidence for this hypothesis has been lacking. Here, we show that naturally occurring amino acid substitutions in the D1 domain of the chimpanzee CD4 not only curb SIVcpz infection, but potentially also guard against cross-species transmission of SIVs infecting monkeys that are hunted by chimpanzees.

**Results**

**Chimpanzee CD4+ T Cell Cultures Differ in Their Susceptibility to SIVcpz Infection.** Generating infectious molecular clones (IMCs) of SIVcpz, 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 SIVcpzPtt (MT145, EK505, MB897, LB715, and GAB2) and SIVcpzPts (BF1167, TAN2, and TAN13) strains (21, 44–47). SIVcpz IMCs were transfected, normalized based on infectivity, and used to infect CD4+ T cells at a multiplicity of infection of 0.1 (Fig. 1). As observed previously (21, 44, 47), all SIVcpz strains replicated efficiently in human CD4+ T cells (Fig. L4). 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, SIVcpz strains (e.g., Melissa in Fig. 1 B 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 SIVcpz 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 SIVcpz was blocked, we generated a replication-competent SIVcpz reporter virus by inserting an enhanced green fluorescent protein (eGFP) gene between the env and nef genes of the SIVcpz 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 SIVcpz was inhibited at the level of cell entry.

**CD4 Polymorphisms Govern the Susceptibility of CD4+ T Cells to SIVcpz Infection.** In addition to CD4, SIVcpz 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. S34). These nonsynonymous SNPs caused substitutions at positions 40 (Q/R), 52 (N/K), 55 (V/I), and 68 (P/T), the combination of which resulted in five CD4 variants that differed solely in their D1 domain (Fig. 1E). Importantly, the P68T substitution created a potential N-linked glycosylation site (PNGS) at position 66, in addition to the invariant PNGS present at position 32 present in A.

**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 phenotype (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 the same set of viruses, except for GAB2. In contrast, chimpanzees homozygous for the QNVT allele supported infection of 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 one or both 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)QONVT 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. S3F). Thus, in contrast to the human CD4, for which the most common D1 variant occurs at a frequency of only $6.6 \times 10^{-5}$ (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. 2A).

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. 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 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 SIVcpzPš strains (Figs. 3 and 4), but CD4 genotyping of 128 chimpanzees from Gombe National Park identified two SIVcpzPš-infected individuals that were R25-heterozygous. Since the virus of one of these individuals (TAN1) has been molecularly cloned (44), we tested its Env in the CD4 transfection assay (Fig. 4A). This analysis showed that TAN1, in contrast to the other SIVcpzPš Envs, was able to utilize the R25 allele (Fig. 4A). Similarly, of five SIVcpzPš 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 (Fig. 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 RNQVP 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 RNQVP allele without altering its ability to utilize the QQNVP allele (Fig. 4E). Similarly, residue 455 tended to be negatively charged in Envs that could use R40-containing CD4 alleles, but neutral in Envs 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 T283, the contact residue for CD4 Q40. However, this loss was predicted to be potentially offset by the formation of long-range salt bridges between the 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 was sufficient to restore its infectivity for QRNVP-expressing cells without altering its ability to utilize the QONVP 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⁺ T cells from a chimpanzee (Chip). But was heterozygous for the QRNVP and QONVP alleles (SI Appendix, Fig. S7A). Although Env amino acid 455 is unlikely the only residue impacted by the R40 substitution (the LP715 Env encodes a T at position 455 but is able to use the QRNVP 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 Envs 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. Immunoprecipitation of CD4 proteins from cells infected with human and chimpanzee (QONVP and QONVT) 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...
with glycans on the various SIVcpz Envs, we modeled their position in the crystal structures 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. 5C). The N460Q mutant also replicated in CD4+ T cells from one, but not all, QQNVT-homozygous chimpanzees (Fig. 5D and E) while N295Q and/or N446Q mutants again had no effect. Thus, removal of an Env glycan predicted to clash with the conserved N32 increased the infectivity of the MB897 strain, both in transient 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.

Fig. 4. SIVcpz cell entry is blocked by charged residues at the CD4–Env interface. (A and C) Percentage of cells expressing the indicated CD4 alleles that are infected with SIVcpz Env-containing pseudoviruses. Points connected by lines represent paired averages of three independent experiments, each performed in triplicate. The parentheses indicate the CD4 genotype (amino acid at positions 25 in A and position 40 in C) of the naturally infected chimpanzee from which the respective SIVcpz Env was derived. (B and D) Modeling of chimpanzee CD4 residues 25 (B) and 40 (D) on the crystal structure of the HIV-1 gp120 (yellow) bound to human CD4 (green). Polymorphic CD4 residues are shown in dark blue (Q) and red (R), respectively. Env residues in proximity to these polymorphic sites are highlighted. Partial Env protein alignments for the tested SIVcpz strains are shown, with the residue predicted to interact with the polymorphic CD4 residue highlighted (blue for Env that can use both Q and R residues, red for Env that is inhibited by the R residue). (E) Infectivity of WT (MB897, TAN2) and mutant (MB897 T455D, TAN2 K474V) SIVcpz Envs on cells expressing the indicated chimpanzee CD4 alleles. Circles represent average infectivity with mean and standard deviation shown. Significant differences are indicated (**P < 0.01, unpaired t test). (F) Replication potential of wild-type (wt) and mutant (T455D) MB897 and wild-type MT145 in chimpanzee CD4+ T cells heterozygous for QRNVP and QQNVT alleles. Viral replication was monitored in culture supernatants by determining reverse transcriptase (RT) activity.

To examine whether removal of the clashing glycans in the MB897 Env would increase its infectivity, we changed the asparagine residues at positions 295, 446, and 460 to glutamines and then tested these mutants, alone and in combination, in the transient CD4 transfection assay (Fig. 5B). Interestingly, the N460Q and the triple mutant increased the infectivity of the MB897 Env in both QRNVP- and QQNVT-expressing cells (Fig. 5B). In contrast, removal of N295 and/or N446 had little effect in the context of either allele (Fig. 5B). Introduction of these same mutations into the MB897 infectious molecular clone were consistent with these results. Compared with WT MB897, both the N460Q and the triple mutant grew faster and to higher titers in CD4+ T cells from a QRNVP/QQNVT-heterozygous chimpanzee (Fig. 5C). The N460Q mutant also replicated in CD4+ T cells from one, but not all, QQNVT-homozygous chimpanzees (Fig. 5D and E) while N295Q and/or N446Q mutants again had no effect. Thus, removal of an Env glycan predicted to clash with the conserved N32 increased the infectivity of the MB897 strain, both in transient 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

![Fig. 4](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

![Fig. 4](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
were further analyzed. Using the transient CD4 transfection assay, we identified nine Envs from SIVs infecting mustached monkeys (Cercopithecus cephus), Hooest's monkeys (Cercopithecus hooesti), western red colobus (Procolobus badius), red-tailed monkeys (Cercopithecus ascanius), sooty mangabeys (Cercocebus atys), and African green monkeys (Cercopithecus aethiops) that mediated entry into cells expressing the human CD4 (Fig. 6d). 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 glycamediated (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 K52 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 Kalande 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 south-eastern 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 whole exon level, the chimpanzee CD4 allele was associated with 3.4-fold lower odds of being infected (P = 0.03, 95% CI: 1.07- to 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. troglodytes and 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
the presence of the given residue (Gombe and (8) Lobéké/Mambéré. 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 (substitution/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 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 encoded at least one RQNVP allele. Thus, one reason for the higher SIVcpz prevalence in the Kalander community, 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/Pts prevalence rates that are particularly 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. 3). 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 chimpanzees (Fig. 1A), at least in part because it lacks ENV glycan predicted to clash with N32 and N66 in the chimpanzee CD4 (Fig. 5A). Indeed, cryo-EM analysis of the MT145 Env trimer 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. 5A), which has the potential to induce unwanted (off-target) antibody responses. Since there are other SIVcpz Envs 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 described (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 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 IMGs 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 CD4 genotype 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 S3 and S6 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**).


