Published as: *Nat Struct Mol Biol*. 2010 May ; 17(5): 608–613.

Structure of a clade C HIV-1 gp120 bound to CD4 and CD4-induced antibody reveals anti-CD4 polyreactivity

Ron Diskin[1], **Paola M. Marcovecchio**[1], and **Pamela J. Bjorkman**[1,2,3]

[1]Division of Biology, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125

[2]Howard Hughes Medical Institute, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125

Abstract

Strategies to combat HIV-1 require structural knowledge of envelope proteins from clade C viruses, the most common in the world. We present the first crystal structure containing a clade C gp120 envelope. The structure, a complex between gp120, the host receptor CD4, and the CD4-induced antibody 21c, reveals that the 21c epitope involves contacts with gp120, a non-self antigen, and with CD4, an auto-antigen. Binding studies using wild-type and mutant CD4 showed that 21c Fab binds CD4 in the absence of gp120, and that binding of 21c to clade C and HIV-2 gp120s requires the crystallographically-observed 21c-CD4 interaction. Additional binding data suggested a role for the gp120 V1V2 loop in creating a high-affinity, but slow-forming, epitope for 21c after CD4 binds. This study represents the first visualization of a potentially autoreactive antibody Fab complexed with both self and non-self antigens.

HIV/AIDS has a disproportionate effect in the developing world, where AIDS has devastated entire countries, especially in sub-Saharan Africa1. Group M HIV-1, which is responsible for the majority of infections worldwide, is divided into 10 clades, of which clade C is the most abundant HIV-1 subtype in Africa and countries with the most infections2. High-resolution structural information concerning the envelope spike that mediates binding to host receptors (CD4 and a chemokine receptor) and fusion of viral and cellular membranes was limited to structures of gp120 monomers derived from North American/Western European clade B viruses3–8 and SIV9. Several of these structures involved ternary complexes between a clade B gp120, soluble CD4 (sCD4), and a CD4-induced (CD4i) antibody3–6.

CD4i antibodies recognize conserved regions of gp120 at or near the binding site for the host co-receptor (the CCR5 or CXCR4 chemokine receptor), which are exposed by a

[3]Correspondence should be addressed to P.J.B. (bjorkman@caltech.edu).

Accession codes. Protein Data Bank: Coordinates and associated structure factors have been deposited with accession codes 3LMJ (21c Fab alone) and 3LQA (CAP210–sCD4–21c complex).

Author contributions

R.D. designed and performed crystallographic and binding experiments, P.M.M. produced Fabs, and P.J.B. oversaw the project.

conformational change caused by CD4 binding10. These antibodies are often cross-reactive but not very potent due to limited steric accessibility when gp120 on the viral membrane is bound to CD4 on the target cell11. However, some CD4i antibodies show reactivity towards free gp120, although less than towards CD4-gp120 complexes, suggesting that unliganded envelope spikes sample different conformations including the CD4-bound conformation10. The 21c CD4i antibody, which shows little binding to free gp120s (R.D. and P.J.B., unpublished results), was isolated from transformed B cells donated by a HIV-1 infected individual classified as a long-term non-progressor12. Although HIV-1 and HIV-2 generally share only weak antigenic cross-reactivity, 21c and other CD4i antibodies found in the sera of HIV-1-infected patients neutralized HIV-2 as well as HIV-1 viruses in the presence of soluble CD4 (sCD4)13.

Here we present the crystal structure of a complex between the clade C gp120 CAP210, sCD4, and the Fab of 21c, and compare the first non-clade B gp120 structure to known clade B gp120 structures3–8. We also describe the potential for autoreactivity by the 21c antibody revealed by the CAP210–sCD4–21c structure, which showed that 21c has a bimolecular epitope involving the expected contacts with gp120 as well as unexpected contacts with sCD4. Binding studies verified the 21c contact with sCD4, showing that 21c bound to sCD4 in the absence of gp120, and that the antibody bound to CAP210, other clade C gp120s, and a HIV-2 gp120 only when the Fab could form the crystallographically-observed contacts with sCD4. To explore the mechanism for the poor clade C neutralization potency of 21c14, we compared binding of the 21c Fab to gp120s including or not including the V1V2 domain, deriving results suggesting that the V1V2 loop generally blocks 21c access to gp120s even in the presence of sCD4. The crystal structure and binding studies presented here suggest an additional mechanism for the CD4 dependence of 21c-like CD4i antibodies – direct contacts with CD4 in a gp120–CD4–CD4i antibody complex – in addition to demonstrating the potential for autoreactivity by 21c and other CD4i antibodies.

RESULTS AND DISCUSSION

Overview of the CAP210–sCD4–21c complex structure

In order to obtain a clade C gp120 structure, we screened 21 clade C gp120s combined with Fabs from six anti-gp120 antibodies and/or sCD4 (domains 1 and 2 from human CD4). gp120s were modified by eliminating conserved but non-essential N-linked glycosylation sites15. In addition, the V1V2 and V3 loops and the N- and C-termini of the proteins were truncated, similar to constructs used for crystallization of clade B gp120s6. Crystals were obtained for several different combinations, most involving a complex between a clade C gp120, sCD4, and the Fab from the CD4i antibody 21c12. Crystals suitable for data collection were obtained from a complex containing the clade C CAP210.2.00.E8 (CAP210)14 gp120, and the structure of CAP210–sCD4–21c was solved to 3.4 Å resolution (R_{cryst} = 23.4%; R_{free} = 32.2%). The structure of the uncomplexed 21c Fab was also solved to 2.2 Å resolution (R_{cryst} = 19.2%; R_{free} = 21.8%) (Supplementary Table 1).

As seen in previously-described clade B gp120–sCD4–CD4i-Fab structures (HXBc2–sCD4–17b, YU2–sCD4–17b, YU2–sCD4–412d, and JR-FL–sCD4–X5)3–6, the CAP210–sCD4–21c structure comprises the globular core gp120, domains 1 and 2 (D1–D2) of CD4, and the

variable heavy (V_H), variable light (V_L), constant heavy (C_H1) and constant light (C_L) domains of a CD4i Fab (Fig. 1). As in the other structures, sCD4 binds to CAP210 gp120 mainly using the C" β-strand of its D1 domain burying ~1100 $Å^2$ of sCD4 surface area, with a gap between the D1 C' strand and the gp120 surface (Fig. 2). Comparison of gp120–sCD4–CD4i-Fab structures shows that CD4i Fabs adopt different orientations, with the light chains of the 21c and X5 Fabs facing sCD4, whereas the heavy chains of the 17b and 412d Fabs are closest to sCD4 (Fig. 1c).

Comparison of clade B and C gp120 structures

The overall structure of the core clade C CAP210 gp120 is similar to the core structures of HXBc2, YU2, and JR-FL clade B gp120s3,4,6: rms deviations of 0.9 Å, 1.1 Å, and 1.0 Å for superpositions of 268, 263, and 266 common carbon-α atoms, respectively, but somewhat more different from clade B structures than they are from each other: rms deviations of 0.7 Å (HXBc2 and YU2; 267 carbon-α atoms), 0.8 Å (HXBc2 and JRFL; 272 carbon-α atoms), and 0.8 Å (YU2 and JRFL; 278 carbon-α atoms) (Fig. 3). CD4-bound gp120s share a heart-shaped structure consisting of the inner and outer domains with a bridging sheet at the base of the heart6 (Fig. 1b). The N- and C-termini of the polypeptide chain are in the inner domain, distal to the bridging sheet. The co-receptor–CD4i Fab binding site is located on the bridging sheet3 and the CD4 binding site is located between the inner and outer domains and restricted by the CD4-binding loop7,16. Five hypervariable regions, V1 – V5, decorate the surface of gp120, forming extended loops6 (Fig. 1b). The loops mask conserved regions that form the CD4 and the chemokine receptor binding sites and, in the case of the V3 loop, determine cell tropism of the virus through interactions with the co-receptor17,18. The first variable region in the CAP210 construct is V4, a 26-residue loop of which 21 residues (Thr392 – Asn412) are disordered. This region was also disordered in the HXBc2 and YU2 structures3,6 (21 and 13 residues, respectively), but ordered in the structure of JR-FL4 (17 residues). The CAP210 V5 region is a 9-residue loop, three residues longer than V5 in clade B gp120s, and it includes two (as compared with one in clade B sequences) potential N-linked glycosylation sites (Supplementary Fig. 1).

The CD4-binding loop, which interacts with D1 of CD4 and forms a central part of the epitope for b12 and other antibodies that target the CD4 binding site7,8,16, differs in conformation between CAP210 and clade B gp120s (Fig. 3). Sequence variations at the base of this loop facilitate resistance to neutralization by b1216. CAP210 Leu363, conserved in all clade C strains, replaces Pro369, found in most clade B gp120s (Supplementary Fig. 1). The Leu363 sidechain extends from the loop in a manner proposed to interfere with b12 binding, consistent with the low potency of b12 in neutralization of CAP210 HIV-116. Another difference is the replacement of Ser364, common in clade B gp120s, with Pro358 in CAP210. This substitution, also in other clade C gp120s, is not correlated with escape from b12 neutralization, but may promote additional flexibility in clade C CD4-binding loops due to loss of a stabilizing intra-loop hydrogen bond in clade B loops between the Ser364 sidechain and the mainchain –NH group of Gly366 (Fig. 3). In common with clade B gp120s, CAP210 Asp362 interacts with sCD4 Arg59, an interaction mimicked by some anti-CD4-binding site antibodies8.

The bimolecular epitope of 21c

Contacts between 21c and CAP210 map exclusively to the antibody heavy chain, as also observed in a b12–HXBc2 complex structure7. The gp120 footprint on 21c involves the H2 and H3 complementarity-determining regions (CDRs) plus a framework loop connecting β-strands D and E (Fig. 4a,b; Supplementary Fig. 2), which interact with the CAP210 bridging sheet and the base of the V1V2 stem, burying a total of 1290 $Å^2$ at the 21c–gp120 interface. Important interactions include Leu105 from the 21c H3 loop anchored in a hydrophobic pocket at the base of V1V2 defined by CAP210 residue Pro123, and a salt bridge between 21c H2 residue Glu55 and CAP210 Arg431. In addition, Glu54 and Asn56 from the 21c H2 loop are in close proximity to two CAP210 arginines (Arg420 and Arg418), likely forming a hydrogen bond network together with solvent molecules (ordered water molecules were not observed due to resolution limitations), and a portion of the 21c framework region near Thr74 buries CAP210 Met433 and Cys118 (Supplementary Fig. 3a).

Although different CD4i antibodies target distinct regions of the chemokine receptor binding site on gp120, 21c is the first CD4i antibody also observed to make direct contacts with CD4 (Fig. 1c) (a total of 750 $Å^2$ of buried surface area at the 21c–sCD4 interface). On sCD4, the interface includes residues from the loops between β-strands A and B and between strands E and F and a short (5–6 residue) helical segment preceding strand D, which deviates from ideal β-strand dihedral angles (Fig. 2a). On 21c, the interface maps mainly to the light chain (L1 and a nearby framework region between 21c β-strands D and E) plus three residues from H3 (Fig. 4a,b). Important contacts include an ionic interaction between 21c L1 residue Lys32 and sCD4 D1 residue Asp53, and framework residue Ser69 contacting sCD4 near Asp10 in a possible hydrogen bond network involving mainchain atoms (Supplementary Fig. 3b). The 21c H3 loop that contacts gp120 also contributes the sCD4 interface: e.g., H3 Leu105 is stabilized in a pocket created by CAP210 and sCD4 Arg59-Ser60, thus contacting both sCD4 and gp120. In addition, the 21c H3 Trp106 sidechain, although disordered, projects toward sCD4 and contributes to the interface, while the mainchain contacts gp120.

Comparison of bound and free 21c structures

The structures of free and complexed 21c Fab were compared by superimposing the variable domains (Supplementary Fig. 4). The most obvious difference is the relative orientation of the V_H-V_L and C_H1-C_L domain pairs. Compared with complexed 21c, the heavy chain of the free Fab adopts a more acute elbow bend angle between the V_H and C_H1 domains and a concomitant increase in the bend angle between the V_L and C_L domains, resulting in a ~31 Å movement of the constant domains. The changes, likely resulting from different crystal packing forces, demonstrate the potential for inter-domain flexibility during binding of the 21c IgG to HIV or virally-infected cells. The largest binding-induced changes are in the H3 and H2 loops, which adopt new conformations that allow Leu105 (H3) and Glu55 (H2) to fit into their respective pockets on gp120.

21c binds sCD4 in the absence of gp120

The contact between sCD4 and the 21c V_L domain in the CAP210–sCD4–21c structure prompted us to investigate whether 21c binds to CD4 in the absence of gp120. For comparison with wild-type sCD4, we constructed a sCD4 mutant ($sCD4_{K75T}$; Lys75 to Thr)

in which a N-linked glycosylation site was introduced in a position predicted to interfere with sCD4 interactions with 21c but not with gp120 (Fig. 5a). SDS-PAGE analysis showed a mobility shift consistent with glycosylation of the introduced Asn–X–Thr motif (Fig. 5b), and a surface plasmon resonance (SPR) binding assay verified that $sCD4_{K75T}$ retained binding to CAP210 gp120 (data not shown). SPR experiments in which increasing concentrations of sCD4 or $sCD4_{K75T}$ were injected over immobilized 21c Fab revealed weak binding for the sCD4-21c interaction (K_D>100 μM), but no detectable binding of $sCD4_{K75T}$ or CAP210 core gp120 to 21c (Fig. 5c). These results demonstrate that 21c Fab binds to sCD4 in the absence of gp120 using the crystallography-observed contacts, and that 21c requires sCD4 for binding to CAP210 gp120.

21c contacts sCD4 in clade C and HIV-2 gp120 complexes

In order to investigate whether the 21c contact with sCD4 is a general characteristic of its interactions with gp120–sCD4 complexes or results from special features related to CAP210, we compared the abilities of sCD4 and $sCD4_{K75T}$ to form gp120–sCD4–21c complexes. The binding of 21c to the CAP210–$sCD4_{K75T}$ complex was greatly reduced as compared to the CAP210–sCD4 complex (Fig. 5d), verifying that the crystallographically-observed 21c contacts with sCD4 occur when 21c binds to CAP210 in solution. Similar results were observed for a comparison of the interactions of sCD4 and $sCD4_{K75T}$ with 21c and the core regions of the two other clade C gp120s (Supplementary Fig. 5), and UC1, a gp120 from a HIV-2 strain that is potently neutralized by 21c plus sCD413 (Fig. 5d).

The roll of V1V2 in masking the 21c epitope

Several conformational changes in gp120 occur after CD4 binding9. One involves the movement of V1V2 to expose the co-receptor binding site and reveal epitopes for CD4i antibodies. Intact IgG forms of CD4i antibodies are often less potent for neutralization than their Fab or scFv counterparts due to steric occlusion of the larger IgGs when the CD4i epitope is exposed in the small volume between the viral and host cell membranes11. Indeed, 21c IgG does not neutralize CAP210 HIV-114. However, the 21c Fab, alone or with added sCD4, also fails to neutralize CAP210 (data not shown), although 21c plus sCD4 binds to core CAP210 gp120. To evaluate the role of variable loops that are missing from the core CAP210 construct in binding to non-truncated envelope spikes, we expressed core CAP210 gp120 containing the V1V2 or V3 loops ($CAP210_{V1V2}$ and $CAP210_{V3}$) and used SPR assays to assay binding of these proteins to immobilized 21c Fab in the presence of sCD4 or $sCD4_{K75T}$ (Fig. 5e). Binding of $CAP210_{V3}$ and core CAP210 were virtually identical (data not shown), suggesting no role for the V3 loop in restricting binding of 21c to CAP210 plus sCD4. Distinct differences were observed, however, for binding of sCD4 plus 21c Fab to $CAP210_{V1V2}$ compared with binding to core CAP210 (Fig. 5e): the on-rate was slower by over three orders of magnitude (k_{on}=2.0 × 10^4 M^{-1}·sec^{-1} versus 2.6 × 10^7 M^{-1}·sec^{-1} for $CAP210_{V1V2}$ and core CAP210, respectively), but the overall affinity was 17-fold higher for 21c Fab plus sCD4 binding to $CAP210_{V1V2}$ (K_D=92 nM) than to core CAP210 (K_D=1.6 μM). The interaction with $CAP210_{V1V2}$ was abrogated using $sCD4_{K75T}$ instead of sCD4 (Fig. 5e). Similar effects were seen when comparing binding of 21c Fab plus sCD4 or $sCD4_{K75T}$ to core UC1 versus a V1V2-containing UC1 ($UC1_{Full\text{-}length}$); i.e., the on-rate for 21c plus sCD4 binding was slower for $UC1_{Full\text{-}length}$ than core UC1, but the

affinity was higher, and addition of sCD4^{K75T} instead of sCD4 diminished the interaction (Fig. 5e). Affinity increases for V1V2-including gp120s implies that 21c is stabilized on gp120 by forming contacts with V1V2. The slower association kinetics for forming a gp120–sCD4–21c complex in the presence of V1V2 are consistent with a model in which there is an equilibrium between multiple V1V2 conformations, with one conformation imparting stabilizing contacts with 21c after CD4 binds and other conformations sterically inhibiting binding.

The 21c bimolecular epitope suggests anti-CD4 autoreactivity

The discovery of an interaction between the 21c Fab and sCD4 suggests the potential for autoreactivity by 21c IgG, correlating with reports that autoantibodies against CD4 are found in 10–15% of HIV-infected individuals19,20. While the 21c epitope on CD4 is located entirely within D1 (Fig. 2), most autoreactive anti-CD4 antibodies target the two membrane-proximal domains of CD4 (D3 and D4)21 (which would be distant from a bound gp120), suggesting that the most common class of autoreactive anti-CD4 antibodies do not also target gp120.

The bimolecular nature of the 21c epitope is reminiscent of the reactivity of other potentially autoreactive anti-HIV antibodies: the 4E10 and 2F5 broadly neutralizing antibodies, which bind to lipids in addition to the membrane proximal region (MPER) of the gp41 subunit of the HIV-1 envelope spike22,23,24. Analogous to our finding that distinct surfaces of the 21c Fab bind to sCD4 and to gp120 (Fig. 4b), 4E10 binds to lipids using one region of its H3 loop and to gp41 using another region of H325. Although anti-MPER antibodies bind only weakly to lipids alone25,26, a recent study demonstrated that 2F5 is sufficiently autoreactive to activate immunological tolerance in a knock-in mouse27, raising the possibility that the weak affinity we observed for 21c Fab binding to sCD4 (Fig. 5c) is also physiologically relevant.

The bimolecular nature of the 21c epitope raises intriguing questions about the origins of 21c-like antibodies, which exhibit both self- and non-self reactivity. One possibility is that the 21c heavy chain was selected to recognize gp120, and the light chain by chance interacted weakly with CD4 in the orientation imposed by the 21c–gp120 interaction. A chance interaction seems unlikely, however, given that one of the L1 residues that contacts sCD4 (Lys32) was induced by somatic mutation from the germline sequence (Supplementary Fig. 2). If self-reactive binding was selected, 21c-like autoreactive antibodies would likely be found only at relatively late stages in a HIV infection, after dysregulation of the host immune system. By analogy to the possibility that anti-MPER antibodies were originally generated by a combination of virion lipids and the HIV-1 envelope spike, the 21c epitope might be created by gp120–CD4 complexes, either generated by free gp120 binding to CD4 on T cells, by CD4 fragments released after necrotic death of infected T cells binding to a virus or released gp120, or by membrane-bound CD4-gp120 complexes on infected T cells.

METHODS

Protein expression, purification, and crystallization

We isolated the 21c heavy and light chain genes and purified the IgG from transiently-transfected HEK293-6E cell supernatants as described28. We made 21c Fab by digesting purified IgG with immobilized papain (Pierce) at 10 mg mL^{-1} and isolated Fab fragments using protein A (GE Healthcare) and Superdex200 16/60 gel filtration chromatography.

We expressed core CAP210 gp120 (truncated as described6) and other gp120s (18 other core clade C gp120s, $CAP210_{V1V2}$, $CAP210_{V3}$, UC1 and $UC1_{Full\text{-}length}$) using baculovirus-infected insect (Hi5) cells. For gp120 expressions, we subcloned genes encoding C-terminally His-tagged gp120s into the pAcGP67B baculovirus transfer vector (BD Biosynthesis) in frame with the gp67 secretion signal. To eliminate seven non-essential N-linked glycosylation sites15, we introduced T89I, N226D, T232I, N285T, S329N, T388I, and N447D substitutions. We purified gp120 proteins from infected cell supernatants by Ni-NTA (Qiagen) and gel filtration chromatography. We used similar methods for expression and purification of C-terminal His-tagged wildtype sCD4 and $sCD4_{K75T}$ (residues 1–186 of mature CD4).

For crystallization trials, we mixed 2 mg of a core gp120 with equimolar amounts of Fab and sCD4 and partially deglycosylated the complex by incubation with 6250 U of Peptide N-Glycosidase F (NEB) for 10 hours at 37°C. We purified PNGase-treated complexes using Superdex200 10/300 gel filtration chromatography. The CAP210–sCD4–21c crystals crystallized in 15%–20% (w/v) PEG monomethyl ether 2000, 0%–5% (v/v) PEG 200, 0.2 M Trimethylamine N-oxide, 0.1 M Tris-HCl pH 8.5 mixed in a 1:1 ratio with the complex at a concentration of OD_{280}=12. The free 21c Fab crystallized in 70% (v/v) 2-methyl-2,4-pentanediol, 0.1 M Hepes pH 7.5 mixed in a 1:1 ratio with 400 μM 21c Fab in a final volume of 400 nl.

Data collection and structure determination

We collected a complete data set to 2.2 Å resolution from cryoprotected 21c Fab crystals at beamline BL12-2 of the Stanford Synchrotron Radiation Lightsource (SSRL) on a MarMosaic-325 CCD detector and processed the data using HKL200029. We found a molecular replacement solution using Phaser30 utilizing two portions of a Fab (V_H-V_L and C_H1-C_L; PDB code 3C0831 as search models. We refined the model using TLS-restrained refinement32 implemented in Refmac33 and used Coot34 for model building. The final model (R_{work}=19.2%; R_{free}=21.8%) includes 3262 protein atoms, 181 water molecules and two 2-methyl-2,4-pentanediol molecules (Supplementary Table 1). 97%, 2.5% and 0.5% of the residues were in the favored, allowed, and disallowed regions, respectively, of the Ramachandran plot. Five residues at the C-terminus of the heavy chain and two residues at the C-terminus of the light chain were disordered.

We collected data to 3.4 Å (~3.9 Å in the direction of b/b*) from cryoprotected crystals of the CAP210–sCD4–21c complex and processed the data without a B-factor correction using MOSFLM35 and Scala36. The crystals usually exhibited weak, highly mosaic and anisotropic diffraction. We solved the structure by molecular replacement using Phaser30 by

searching for four components; V_H-V_L and C_H1-C_L from the 21c Fab structure, the HXBc2 core plus CD4 D1 (PDB code 2NXY7, and CD4 D2. We performed ellipsoid truncation and anisotropic scaling using the Diffraction Anisotropy Server37, resulting in greatly improved electron density maps and data that were 79% complete to 3.4 Å. We refined the model using CNS38,39, group B factors, and composite omit maps to R_{work}=23.0%; R_{free}=32.1% (R_{free} based on 5% of the reflections). To finalize the model we performed an additional refinement step using a reduced set of reflections for calculating R_{free} (0.5%) in a maximum log likelihood function refinement.

The final model includes 6894 total protein atoms (2272 atoms of CAP210, 1362 atoms of sCD4, and 3260 atoms of Fab 21c), and 56 carbohydrate atoms in four ordered N-acetylglucosamine molecules (attached to gp120 Asn441, Asn380, Asn272 and Asn258). Some residues (mainly surface-exposed) were missing portions of sidechain electron density. We assigned potential aspartic acid residues at N-linked glycosylation sites (formed from asparagines by treatment with PNGase F) as asparagines. The electron density at some solvent-exposed regions was poor, but was higher quality at protein interfaces (Supplementary Fig. 3a). The 21c Fab heavy chain portion of the model includes the same number of residues as in the free 21c structure, and the light chain includes all but the N- and C-terminal residues. sCD4 is missing one residue at the N-terminus and 16 residues at the C-terminus including the 6x-His tag. Disordered residues in the CAP210 structure are indicated in Supplementary Fig. 1. 97.0%, 1.3% and 1.7% of the residues were in the favored, allowed, and disallowed regions, respectively, of the Ramachandran plot.

We used AreaIMol in CCP436 and a 1.4 Å probe for calculated buried surface areas and PyMol40 for superposition calculations and molecular representations.

SPR binding experiments

We evaluated interactions between 21c Fab, gp120, and sCD4 using a Biacore T100 instrument (GE Healthcare). Purified 21c Fab (0.5 μM) was covalently immobilized to three of the four flow cells on a CM5 sensor chip (Biacore) in 10 mM acetate pH 5.5 using standard primary amine coupling chemistry at three different densities: 250, 500, and 1000 resonance units (RU). The fourth flow cell was mock coupled using buffer to serve as a blank. We performed measurements at 25°C in 20 mM Hepes pH 7.0, 150 mM NaCl, and 0.005% surfactant P20, and regenerated sensor chips using 10 mM glycine pH 2.5. In some experiments, we injected a dilution series of sCD4, $sCD4_{K75T}$, or CAP210 gp120 and reactions were allowed to closely approach or reach equilibrium. We derived a lower estimate for the equilibrium dissociation constant (K_D) for the 21c-sCD4 interaction by nonlinear regression analysis of a plot of R_{eq} (the equilibrium binding response) versus the log injected protein concentration using a 1:1 binding model (Fig. 5c). In other experiments (Fig. 5d,e), we injected a dilution series of a gp120 incubated with 2 μM sCD4 or $sCD4_{K75T}$ at a flow rate of 90 μl min^{-1}, and monitored the association and dissociation phases to derive kinetic constants. After subtracting the signal from the mock-coupled flow cell, we globally fit the kinetic data to a 1:1 binding model (Biacore evaluation software) to derive on- and off-rate constants, which were used to calculate the affinity as $K_D=k_{off}/k_{on}$. We confirmed the low affinity for the CAP210 plus sCD4 interaction with 21c Fab (K_D=1.6 μM; Fig. 5c)

by nonlinear regression analysis of a R_{eq} versus log injected protein concentration plot in which the highest concentration of injected CAP210 was 10 μM (data not shown).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by a fellowship from the European Molecular Biology Organization (EMBO) (R.D.), a Collaboration for AIDS Vaccine Discovery (CAVD) grant with support from the Bill & Melinda Gates Foundation (Grant 38660; PI: L. Stamatatos) (P.J.B.), and the Molecular Observatory at Caltech supported by the Gordon and Betty Moore Foundation. Operations at SSRL are supported by the US-DOE and NIH. We thank Roland Strong for suggesting the sCD4 mutant experiments, Inderjit Nangiana and the Caltech Protein Expression Center for expressing proteins in insect cells, Yunji Wu for testing for tyrosine sulfation on 21c, Rachel Galimidi and Maria Politzer for technical assistance, Marta Murphy for help with figures and David Baltimore, Leo Stamatatos, and Anthony West for critical reading of the manuscript.

References and Notes

1. UNAIDS. Report on the global HIV/AIDS epidemic. 2008.
2. McCutchan FE. Understanding the genetic diversity of HIV-1. AIDS. 2000; 14(Suppl 3):S31–S44. [PubMed: 11086847]
3. Huang CC, et al. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. Science. 2007; 317:1930–1934. [PubMed: 17901336]
4. Huang CC, et al. Structure of a V3-containing HIV-1 gp120 core. Science. 2005; 310:1025–1028. [PubMed: 16284180]
5. Kwong PD, et al. Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. Structure. 2000; 8:1329–1339. [PubMed: 11188697]
6. Kwong PD, et al. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature. 1998; 393:648–659. [PubMed: 9641677]
7. Zhou T, et al. Structural definition of a conserved neutralization epitope on HIV-1 gp120. Nature. 2007; 445:732–737. [PubMed: 17301785]
8. Chen L, et al. Structural Basis of Immune Evasion at the Site of CD4 Attachment on HIV-1 gp120. Science. 2009; 326:1123–1127. [PubMed: 19965434]
9. Chen B, et al. Structure of an unliganded simian immunodeficiency virus gp120 core. Nature. 2005; 433:834–841. [PubMed: 15729334]
10. DeVico AL. CD4-induced epitopes in the HIV envelope glycoprotein, gp120. Curr HIV Res. 2007; 5:561–571. [PubMed: 18045112]
11. Labrijn AF, et al. Access of antibody molecules to the conserved coreceptor binding site on glycoprotein gp120 is sterically restricted on primary human immunodeficiency virus type 1. J Virol. 2003; 77:10557–10565. [PubMed: 12970440]
12. Xiang SH, Doka N, Choudhary RK, Sodroski J, Robinson JE. Characterization of CD4-induced epitopes on the HIV type 1 gp120 envelope glycoprotein recognized by neutralizing human monoclonal antibodies. AIDS Res Hum Retroviruses. 2002; 18:1207–1217. [PubMed: 12487827]
13. Decker JM, et al. Antigenic conservation and immunogenicity of the HIV coreceptor binding site. J Exp Med. 2005; 201:1407–1419. [PubMed: 15867093]
14. Li M, et al. Genetic and neutralization properties of subtype C human immunodeficiency virus type 1 molecular env clones from acute and early heterosexually acquired infections in Southern Africa. J Virol. 2006; 80:11776–11790. [PubMed: 16971434]
15. Balzarini J, et al. Carbohydrate-binding agents cause deletions of highly conserved glycosylation sites in HIV GP120: a new therapeutic concept to hit the achilles heel of HIV. J Biol Chem. 2005; 280:41005–41014. [PubMed: 16183648]

16. Wu X, et al. Mechanism of human immunodeficiency virus type 1 resistance to monoclonal antibody B12 that effectively targets the site of CD4 attachment. J Virol. 2009; 83:10892–10907. [PubMed: 19692465]
17. Saunders CJ, et al. The V1, V2, and V3 regions of the human immunodeficiency virus type 1 envelope differentially affect the viral phenotype in an isolate-dependent manner. J Virol. 2005; 79:9069–9080. [PubMed: 15994801]
18. Hartley O, Klasse PJ, Sattentau QJ, Moore JP. V3: HIV's switch-hitter. AIDS Res Hum Retroviruses. 2005; 21:171–189. [PubMed: 15725757]
19. Kowalski M, et al. Antibodies to CD4 in individuals infected with human immunodeficiency virus type 1. Proc Natl Acad Sci U S A. 1989; 86:3346–3350. [PubMed: 2541442]
20. Chams V, Jouault T, Fenouillet E, Gluckman JC, Klatzmann D. Detection of anti-CD4 autoantibodies in the sera of HIV-infected patients using recombinant soluble CD4 molecules. AIDS. 1988; 2:353–361. [PubMed: 3146263]
21. Denisova G, et al. Characterization of new monoclonal antibodies that discriminate between soluble and membrane CD4 and compete with human anti-CD4 autoimmune sera. Mol Immunol. 2003; 40:231–239. [PubMed: 12943795]
22. Haynes BF, et al. Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. Science. 2005; 308:1906–1908. [PubMed: 15860590]
23. Cardoso RM, et al. Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. Immunity. 2005; 22:163–173. [PubMed: 15723805]
24. Ofek G, et al. Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. J Virol. 2004; 78:10724–10737. [PubMed: 15367639]
25. Alam SM, et al. Role of HIV membrane in neutralization by two broadly neutralizing antibodies. Proc Natl Acad Sci U S A. 2009; 106:20234–20239. [PubMed: 19906992]
26. Xu H, et al. Interactions between Lipids and Human Anti-HIV Antibody 4E10 Can Be Reduced without Ablating Neutralizing Activity. J. Virol. 2010; 84:1076–1088. [PubMed: 19906921]
27. Verkoczy L, et al. Autoreactivity in an HIV-1 broadly reactive neutralizing antibody variable region heavy chain induces immunologic tolerance. Proc Natl Acad Sci U S A. 2010; 107:181–186. [PubMed: 20018688]
28. West AP Jr, et al. Evaluation of CD4-CD4i antibody architectures yields potent, broadly cross-reactive anti-HIV reagents. J. Virol. 2009 JVI.01528-01509.
29. Otwinowski, Z.; Minor, W. Methods in Enzymology. Vol. Vol. 276. Academic Press; 1997. Processing of X-ray Diffraction Data Collected in Oscillation Mode; p. 307-326.
30. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
31. Schmiedel J, Blaukat A, Li S, Knochel T, Ferguson KM. Matuzumab binding to EGFR prevents the conformational rearrangement required for dimerization. Cancer Cell. 2008; 13:365–373. [PubMed: 18394559]
32. Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr D Biol Crystallogr. 2006; 62:439–450. [PubMed: 16552146]
33. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1997; 53:240–255. [PubMed: 15299926]
34. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]
35. Leslie AGW. Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography. 1992; 26
36. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr. 1994; 50:760–763. [PubMed: 15299374]

37. Strong M, et al. Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. Proc Natl Acad Sci U S A. 2006; 103:8060–8065. [PubMed: 16690741]

38. Brunger AT, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr. 1998; 54:905–921. [PubMed: 9757107]

39. Brunger AT. Version 1.2 of the Crystallography and NMR system. Nat Protoc. 2007; 2:2728–2733. [PubMed: 18007608]

40. DeLano, WL. The PyMOL Molecular Graphics System. Palo Alto: DeLano Scientific; 2002.

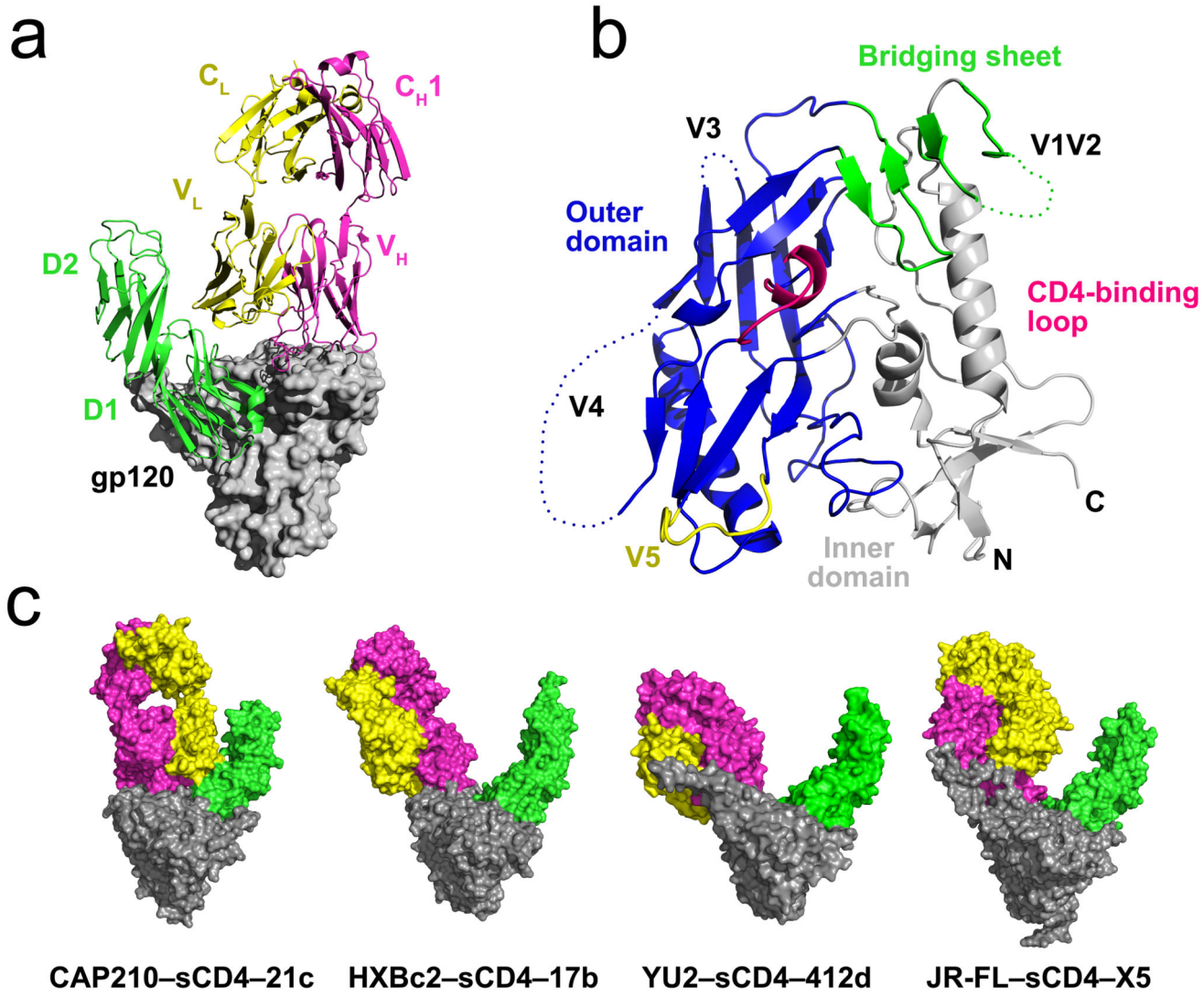

Fig. 1.

Structure of CAP210–sCD4–21c and comparison with other gp120–sCD4–CD4i-Fab complexes. (a) Ribbon representations of the 21c Fab (heavy chain in pink; light chain in yellow) and sCD4 (green) complexed with the clade C CAP210 gp120 (gray surface). (b) Ribbon representation of the CAP210 core gp120 structure with variable loops and other features highlighted. (c) Comparison of the CAP210–sCD4–21c structure with structures of sCD4–CD4i-Fab complexes including clade B gp120s (PDB codes 1G9M[5], 2QAD[3], 2B4C[4] for HXBc2–sCD4–17b, YU2–sCD4–412d, and JR-FL–sCD4–X5, respectively). Structures are shown in surface representation with the Fab heavy chain in pink, the light chain in yellow, sCD4 in green and gp120 in gray.

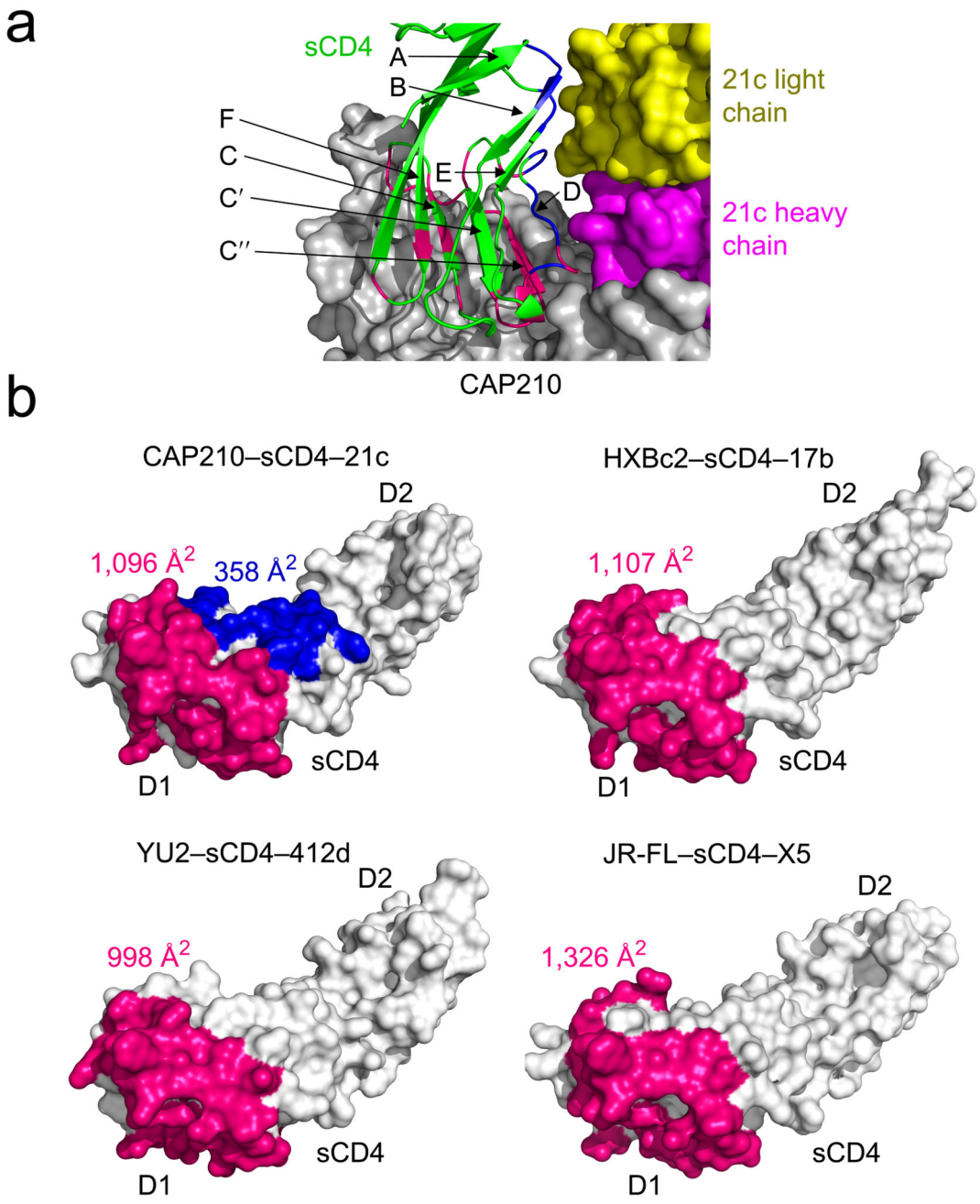

Fig. 2.

sCD4 interactions. (a) Close-up of the interaction between sCD4 D1 (green ribbons with letters referring to individual β-strands, blue highlights for 21c-contacting regions, and pink highlights for CAP210-contacting regions), CAP210 gp120 (gray), and the 21c heavy and light chains (pink and yellow, respectively). The sCD4 C' strand is near the interface with gp120, but forms no contacts within 1.4 Å, creating a gap at the sCD4–gp120 interface as seen in previous structures3,4,6. (b) Surface representations of sCD4 from different gp120–sCD4–CD4i-Fab structures (PDB codes: HXBc2–sCD4–17b: 1G9M[5]; YU2–sCD4–412d:

2QAD[3]; JR-FL–sCD4–X5: 2B4C[4]). sCD4 surface area that is buried at the interface with gp120 is highlighted in pink on each sCD4 structure with the number of Å^2 buried on sCD4 by gp120 indicated. The CAP210–sCD4–21c structure is the only complex structure in which the CD4i Fab also contacts sCD4 (blue highlighted area along with the number of Å^2 buried on sCD4 by 21c).

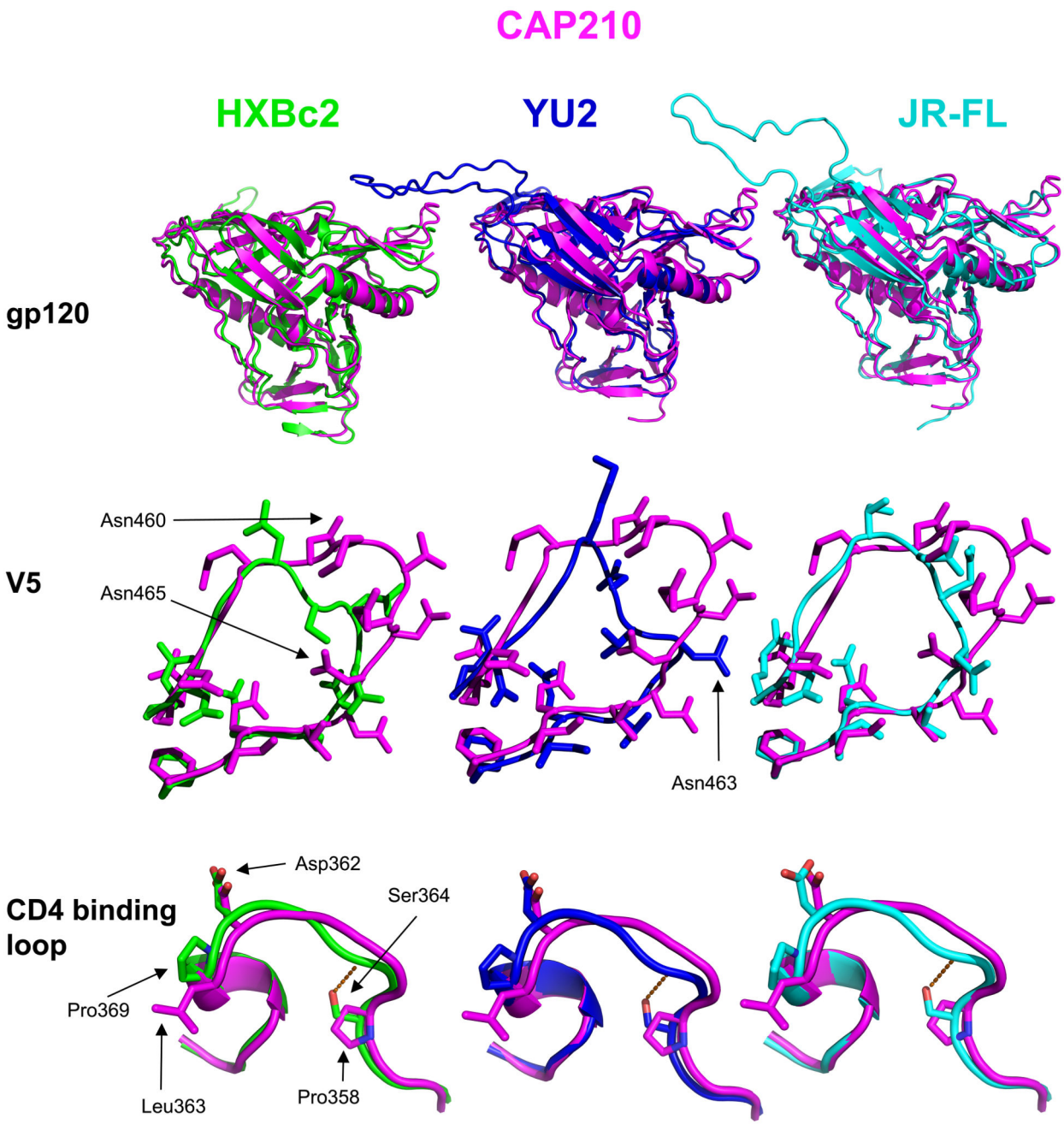

Fig. 3.

Comparison of CAP210 gp120 with clade B gp120 structures. The core CAP210 structure was superimposed with the HXBc2, YU2, and JR-FL structures (PDB codes 1G9M[5], 2QAD[3], 2B4C[4] for HXBc2–sCD4–17b, YU2–sCD4–412d, and JR-FL–sCD4–X5, respectively) (top). The middle and bottom panels show superpositions of the V5 and CD4-binding loops. In the CAP210 V5 loop, the conserved N-linked site at Asn465 (comparable to Asn463 of YU2 to which carbohydrate is attached3,5) may not be glycosylated as the asparagine sidechain points to the interior of the V5 loop. The longer length of the V5 loop

in CAP210 and other clade C gp120s as compared with clade B gp120s (Supplementary Fig. 1) may restrict access to the CD4 binding site.

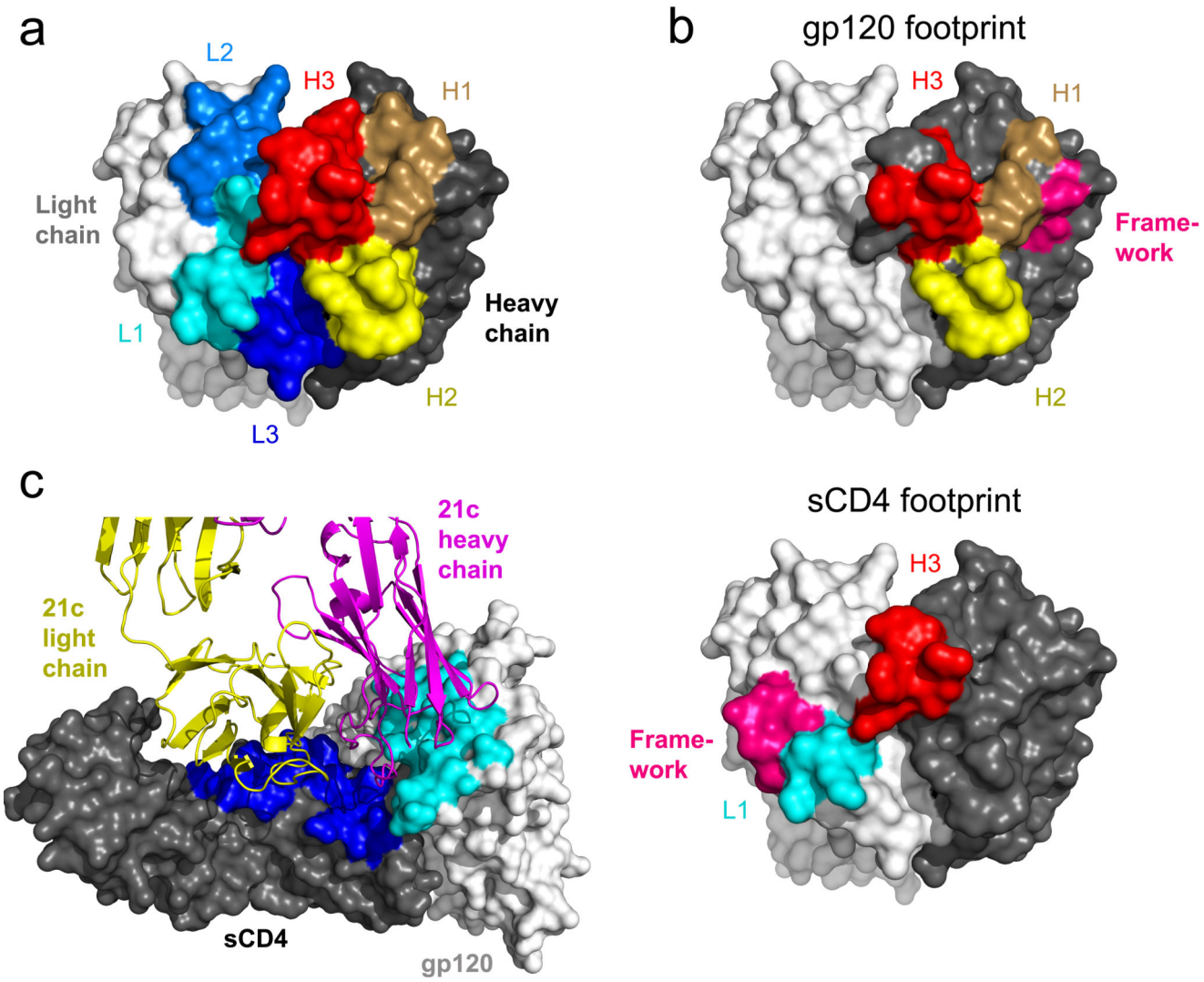

Fig. 4.

Contact surfaces on 21c, sCD4, and gp120. (a) Surface representation of the 21c combining site (from the CAP210–sCD4–21c structure) with CDR1s in gold (H1) and cyan (L1); CDR2s in yellow (H2) and marine blue (L2); and CDR3s in red (H3) and dark blue (L3). (b) Footprints of gp120 (top) and sCD4 (bottom) on the 21c combining site. CDR contacts highlighted as in (a) with framework region contacts in pink. (c) 21c footprint on sCD4 (dark blue) and CAP210 (cyan).

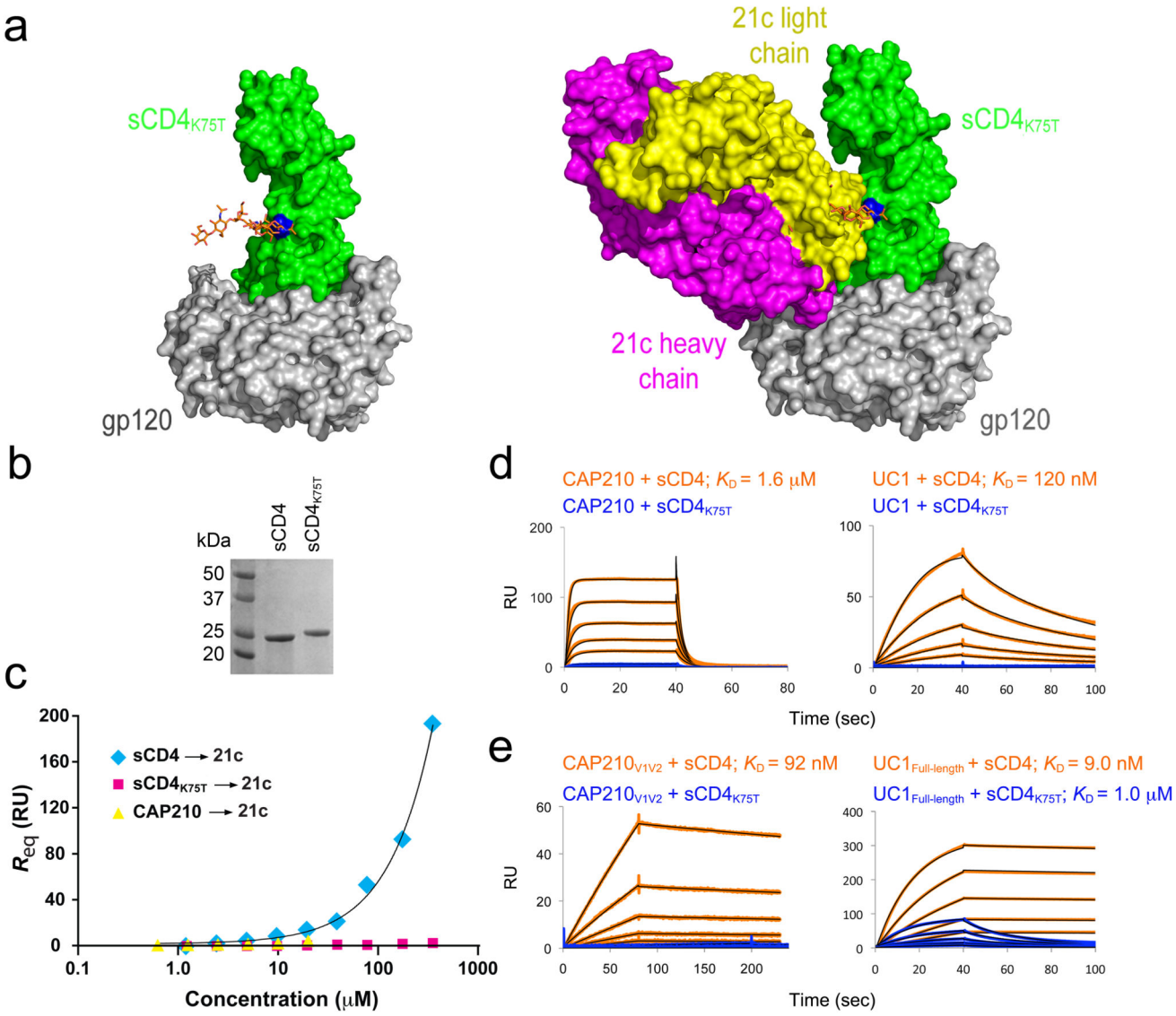

Fig. 5.

21c interactions with sCD4 and gp120s. (a) Surface representation of CAP210–sCD4 (left) and CAP210–sCD4–21c (right) with a modeled carbohydrate at the location of the introduced N-linked glycosylation site (blue) in $sCD4_{K75T}$. The introduced glycan is predicted to disrupt interactions with 21c, but is distant from the interface with gp120. (b) SDS-PAGE comparison of the mobility of wild-type sCD4 and $sCD4_{K75T}$. The mobility shift to a higher apparent molecular weight of $sCD4_{K75T}$ compared with sCD4 is consistent with glycosylation at the introduced N-linked glycosylation site. (c) Equilibrium binding data (equilibrium binding response (R_{eq}) versus the log of the indicated protein concentrations) for SPR experiments in which sCD4, the $sCD4_{K75T}$ mutant, or CAP210 gp120 were injected over immobilized 21c. The best-fit binding curve to the experimental data is shown for the sCD4–21c interaction. (d) Sensorgrams from SPR experiments in which a concentration series (from 1.0 μM–31.2 nM, two-fold dilutions) of the indicated

core gp120 protein (clade C CAP210 or HIV-2 UC1) was injected together with 2 μM sCD4 (orange curves) or $sCD4_{K75T}$ (blue curves, thickened for clarity) over immobilized 21c Fab (reproduced with residual plots in Supplementary Fig. 5). (e) Sensorgrams from SPR experiments in which a concentration series (from 1.0 μM–31.2 nM, two-fold dilutions) of the indicated gp120 protein containing the V1V2 loop ($CAP210_{V1V2}$ or $UC1_{Full\text{-}length}$) was injected together with 2 μM sCD4 (orange curves) or $sCD4_{K75T}$ (blue curves) over immobilized 21c Fab (reproduced with residual plots in Supplementary Fig. 5).