

# Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site

Harry B Gristick<sup>1,9</sup>, Lotta von Boehmer<sup>2,9</sup>, Anthony P West Jr<sup>1</sup>, Michael Schamber<sup>1</sup>, Anna Gazumyan<sup>2</sup>, Jovana Goljanin<sup>2</sup>, Michael S Seaman<sup>3</sup>, Gerd Fätkenheuer<sup>4,5</sup>, Florian Klein<sup>5-7</sup>, Michel C Nussenzweig<sup>2,8</sup> & Pamela J Bjorkman<sup>1</sup>

**HIV-1 vaccine design is informed by structural studies elucidating mechanisms by which broadly neutralizing antibodies (bNAbs) recognize and/or accommodate N-glycans on the trimeric envelope glycoprotein (Env). Variability in high-mannose and complex-type Env glycoforms leads to heterogeneity that usually precludes visualization of the native glycan shield.**

We present 3.5-Å- and 3.9-Å-resolution crystal structures of the HIV-1 Env trimer with fully processed and native glycosylation, revealing a glycan shield of high-mannose and complex-type N-glycans, which we used to define complete epitopes of two bNAbs. Env trimer was complexed with 10-1074 (against the V3-loop) and IOMA, a new CD4-binding site (CD4bs) antibody. Although IOMA derives from VH1-2\*02, the germline gene of CD4bs-targeting VRC01-class bNAbs, its light chain lacks the short CDRL3 that defines VRC01-class bNAbs. Thus IOMA resembles 8ANC131-class/VH1-46-derived CD4bs bNAbs, which have normal-length CDRL3s. The existence of bNAbs that combine features of VRC01-class and 8ANC131-class antibodies has implications for immunization strategies targeting VRC01-like bNAbs.

HIV-1 Env, the only target of neutralizing antibodies, is among the most heavily glycosylated proteins ever characterized<sup>1</sup>. It includes glycans, constituting up to 50% of its mass, attached to  $30 \pm 3$  potential N-linked glycosylation sites (PNGSs) per gp120–gp41 protomer. Viral glycans are generally nonimmunogenic because they are assembled by host-cell machinery; thus, carbohydrates decorating the surface of Env constitute a ‘glycan shield’ that reduces access to underlying protein epitopes<sup>2</sup>. Structural studies of bNAbs bound to Env trimers have revealed mechanisms by which bNAbs targeting various epitopes penetrate the glycan shield to either accommodate or include N-glycans in their epitopes<sup>3–9</sup>. The structure of the HIV-1 Env glycan shield itself, however, remains incompletely characterized.

Because of steric constraints that limit the activities of endoplasmic reticulum and Golgi carbohydrate-processing enzymes, the HIV-1 Env glycoprotein includes regions of underprocessed N-glycans in oligomannose forms ( $\text{Man}_{5-9}\text{GlcNAc}_2$ ), particularly in the intrinsic mannose patch on gp120, which forms portions of the epitopes for many characterized HIV-1 bNAbs<sup>10</sup>. Although oligomannose glycans dominate parts of HIV-1 Env, such as the  $\text{N}332_{\text{gp}120}$  glycan-associated region on gp120, processed complex-type N-glycans predominate at N-linked glycosylation sites on gp41 and gp41-proximal regions of gp120 (ref. 11) and are thought to protect the host receptor (CD4)-binding site (CD4bs) and the V3 loop of gp120 (ref. 12).

Because all Env crystal structures to date have been solved by using glycoproteins produced in exclusively high-mannose forms<sup>2–6,9,13–20</sup>,

little is known about the structure of the native HIV-1 Env glycan shield that includes both complex-type and oligomannose N-glycans. A 4.2-Å cryo-EM structure of a natively glycosylated Env trimer reveals some ordered complex-type N-glycans near the gp120–gp41 interface, but much of the remaining Env glycan content is not visible<sup>8</sup>. Thus, the natively glycosylated epitopes of important HIV-1 bNAbs, such as  $\text{N}332_{\text{gp}120}$  glycan/V3 loop and CD4bs bNAbs, remain incompletely characterized. Here we set out to define the complete epitopes of two HIV-1 bNAbs, which recognize the gp120 V3 loop and CD4bs, respectively, in the context of a natively glycosylated Env trimer.

The accessibility to the conserved CD4bs on gp120 is restricted by surrounding glycans that have, to date, been visualized in structural studies involving Env proteins including only high-mannose N-glycans<sup>9,13,14,18–20</sup>. However, vaccine-design efforts to elicit CD4-mimetic bNAbs<sup>21–27</sup> have been aided by crystal structures revealing that VRC01-class bNAbs mimic CD4 binding and thus share a common mode of gp120 binding and glycan accommodation by using a VH1-2\*02-derived variable heavy ( $V_H$ ) domain<sup>14,18–20,28</sup>. Although VRC01-class bNAbs are attractive candidates for immunogen design, their features, such as a high degree of somatic hypermutation (SHM) and a short (five-residue) light chain (LC) complementarity-determining region 3 (CDRL3) (found in only 1% of human LCs<sup>29</sup>) suggest that they might be difficult to elicit through vaccination. Here we describe IOMA, a new class of CD4-mimetic bNAb derived from the

<sup>1</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California, USA. <sup>2</sup>Laboratory of Molecular Immunology, The Rockefeller University, New York, New York, USA. <sup>3</sup>Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA. <sup>4</sup>Department of Internal Medicine I, University Hospital of Cologne, Cologne, Germany. <sup>5</sup>German Center for Infection Research (DZIF), Partner Site Bonn-Cologne, Cologne, Germany. <sup>6</sup>Laboratory of Experimental Immunology, Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany. <sup>7</sup>Department of Internal Medicine I, Center of Integrated Oncology Cologne-Bonn, University Hospital Cologne, Cologne, Germany. <sup>8</sup>Howard Hughes Medical Institute, The Rockefeller University, New York, New York, USA. <sup>9</sup>These authors contributed equally to this work. Correspondence should be addressed to P.J.B. (bjorkman@caltech.edu).

Received 12 July; accepted 10 August; published online 12 September 2016; doi:10.1038/nsmb.3291

VH1-2\*02 germline, which has a normal-length CDRL3 and includes fewer SHMs than do VRC01-class bNAbs. We defined IOMA's complete epitope, and the epitope of the V3-loop-directed bNAb 10-1074 (ref. 30), by using crystal structures of a fully and natively glycosylated Env trimer. Analysis of the native glycan shield on HIV-1 Env allowed us to provide what is, to our knowledge, the first full description of the interplay between heterogeneous untrimmed high-mannose and complex-type N-glycans within the CD4bs and V3-loop epitopes on Env, thereby revealing antibody-vulnerable glycan holes and roles of complex-type N-glycans on Env that are relevant to vaccine design.

## RESULTS

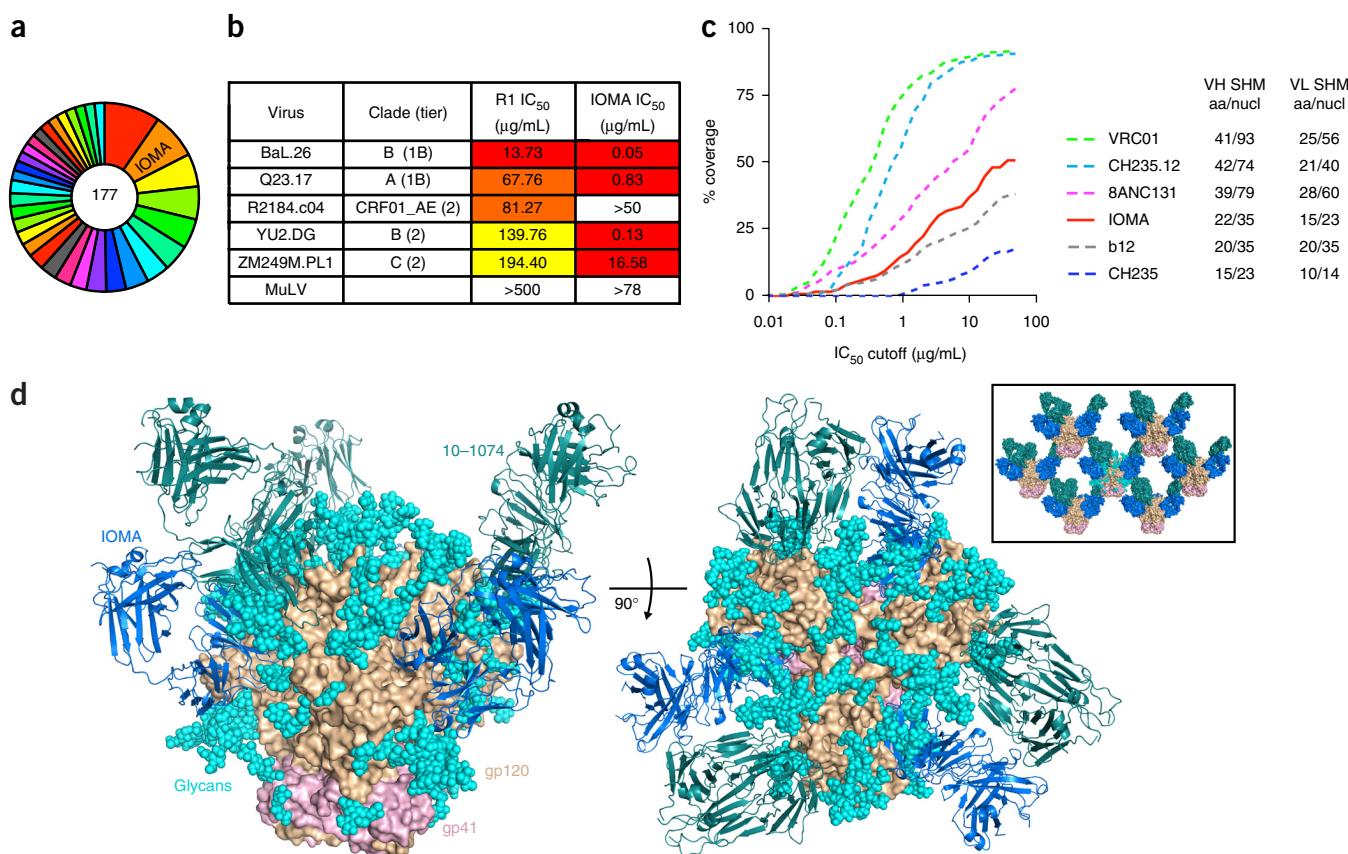
### Isolation and characterization of IOMA

IOMA, a new VH1-2-derived CD4bs bNAb (Fig. 1 and Supplementary Figs. 1–3), was isolated by single-cell B-cell cloning<sup>31</sup> from an HIV-1-infected ART-treated patient. IOMA accounts for most of the neutralizing activity in this patient's serum, as demonstrated by a comparison of neutralization activities of serum IgG and purified IOMA (Fig. 1b). Unlike other CD4bs bNAbs, IOMA has relatively few SHMs (22 HC and 15 LC amino acid changes from the germline, as compared with 41 and 25 changes for VRC01), yet IOMA neutralized ~50% of tested strains with a mean half-maximal inhibitory concentration ( $IC_{50}$ ) of 2.3  $\mu$ g/mL (Supplementary Fig. 2) and a breadth and

potency superior to those of CD4bs bNAbs with similar SHM rates (for example, b12) but inferior to those of CD4bs bNAbs with increased SHMs<sup>32,33</sup> (Fig. 1c). IOMA is an interesting target for structural studies because it combines features of VH1-2/VRC01-class bNAbs with a second related but distinct set of CD4-mimetic CD4bs bNAbs derived from the VH1-46 germline gene segment<sup>19</sup>. Like VH1-46, but unlike VH1-2 bNAbs, IOMA includes a normal-length (eight residue) CDRL3 (Supplementary Note and Supplementary Fig. 3).

### Structures of natively glycosylated Env-bNAb complexes

Crystallization trials were conducted with complexes of the native-like Env trimer BG505 SOSIP.664 (ref. 34) and bNAb Fabs. Previous trimer crystal structures have included Env produced in cells that attached only high-mannose-type N-glycans<sup>3–6,9</sup>, which, with two exceptions<sup>5,9</sup>, were enzymatically trimmed to reduce glycans to single *N*-acetylglucosamines (GlcNAcs) at accessible PNGSs. Our crystals, obtained from natively glycosylated BG505 SOSIP.664 prepared from human cells that attached both complex-type and high-mannose N-glycans, were complexed with Fabs from IOMA and 10-1074, a V3 loop/N332<sub>gp120</sub> glycan-directed bNAb in the PGT121-124 family<sup>30</sup>. We solved independent structures of the IOMA-10-1074-BG505 complex by using BG505 protein prepared from different size-exclusion chromatography (SEC) fractions (Supplementary Fig. 4)



**Figure 1** Isolation and characterization of IOMA. (a) Pie chart showing IgG HC clones of single-cell-sorted HIV-1-antigen-specific memory B cells from patient R1. The number in the center is the number of sequences considered: each colored slice represents one clone, and the slice size is proportional to the number of clonal sequences. (b) Comparison of neutralizing activity of subject-R1 serum IgG and IOMA, an antibody cloned from the second-largest clone of the IgG+ sorted memory B cells (a). (c) Neutralization coverage curves for selected CD4bs bNAbs. The number of SHMs (including indels) at the level of amino acids (aa) and nucleotides (nucl) are indicated for VH and VL gene segments. CH235 is an earlier member of the lineage that produced CH235.12 (ref. 32). (d) Side and top views of the IOMA-10-1074-BG505 structure. Fabs are shown as ribbons, BG505 is shown in surface representation, and glycans are shown as cyan spheres. Inset, packing in IOMA-10-1074-BG505 crystals, demonstrating that the lattice is formed through interactions between neighboring Fabs.

**Table 1** Data collection and refinement statistics

	IOMA-10-1074-BG505 (3.9 Å) <sup>a</sup> fractions 7 and 8 <sup>c</sup> PDB 5T3X	IOMA-10-1074-BG505 (3.5 Å) <sup>b</sup> fractions 11 and 12 <sup>c</sup> PDB 5T3Z
<b>Data collection</b>		
Space group	R3:H	R3:H
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	217.51, 217.51, 156.01	217.26, 217.26, 154.93
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 120	90, 90, 120
Resolution (Å)	80.50–3.90 (4.17–3.90) <sup>d</sup>	80.41–3.50 (3.67–3.50) <sup>d</sup>
$R_{\text{merge}}$	0.35 (2.85)	0.68 (9.29)
$R_{\text{pim}}$	0.08 (0.68)	0.09 (1.21)
$I/\sigma(I)$	7.6 (1.5)	12.6 (1.6)
$CC_{1/2}$	0.99 (0.30)	0.99 (0.40)
Completeness (%)	100 (100)	100 (100)
Redundancy	18.2 (18.3)	61.1 (59.7)
<b>Refinement</b>		
Resolution (Å)	72.07–3.90	51.64–3.50
No. reflections	25,057	34,359
$R_{\text{work}} / R_{\text{free}}$	0.295 / 0.335	0.273 / 0.295
No. atoms		
Protein	11,208	11,212
Ligand/ion	1,427	1,098
<i>B</i> factors		
Protein	267.26	207.37
Ligand/ion	291.24	225.72
R.m.s. deviations		
Bond lengths (Å)	0.004	0.005
Bond angles (°)	1.17	1.23

<sup>a</sup>The 3.9-Å data set includes four crystals. <sup>b</sup>The 3.5-Å data set includes 14 crystals.<sup>c</sup>Definition of fractions from SEC column are in **Supplementary Figure 4**; *n* = number of crystals used for each structure during data processing. <sup>d</sup>Values in parentheses are for the highest-resolution shell.

at resolutions of 3.5 Å and 3.9 Å (**Table 1**). The IOMA-10-1074-BG505 structures revealed an Env trimer bound to three 10-1074 and three IOMA Fabs (**Fig. 1d**). 19 N-glycans (one GlcNAc up to complex-type tetra-antennary) were visible per gp120-gp41 protomer, forming arrays of glycans extending ~30 Å from the trimer surface (**Fig. 2a–c**). We observed apparent N-glycan differences in the Env portions of the 3.5-Å- and 3.9-Å-resolution structures (**Fig. 2a** and **Supplementary Figs. 5–8**).

### Glycan interpretation and refinement

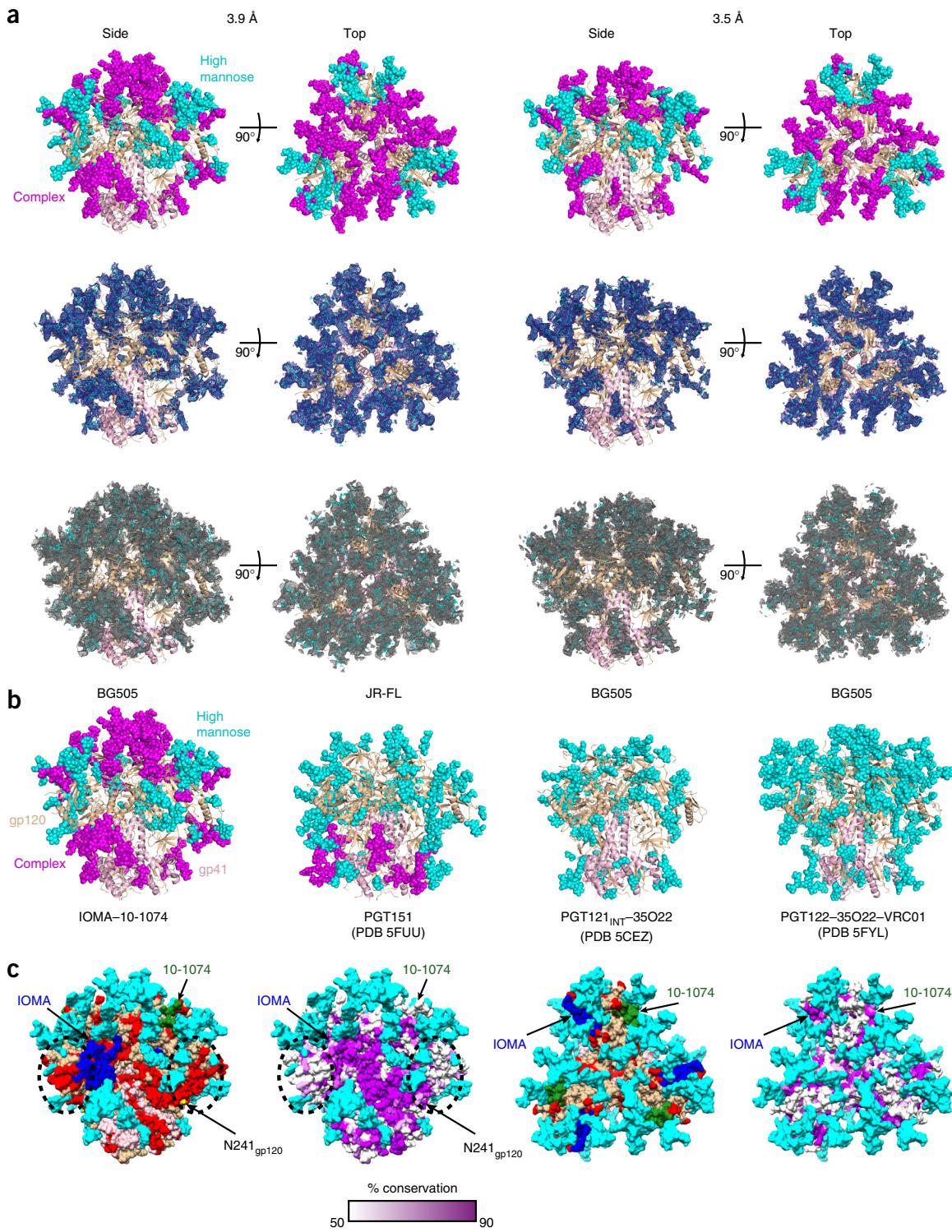
We interpreted glycans in the 3.9-Å and 3.5-Å IOMA-10-1074-BG505 structures by using  $2F_o - F_c$  maps calculated with model phases and using composite annealed omit maps calculated with phases in which the model was omitted to reduce model bias<sup>35</sup>. Although glycan heterogeneity complicated modeling, we were often able to assign glycans at individual PNGSs as complex type or high mannose (**Supplementary Figs. 5–8**). The 3.9-Å structure sometimes showed more density for individual BG505 N-glycans than the 3.5-Å structure (**Supplementary Figs. 5–8**), a result consistent with the apparently greater degree of glycosylation suggested by the SEC and SDS-PAGE characteristics of the BG505 protein in crystals used for the lower-resolution structure (**Supplementary Fig. 4**). We assigned glycans as complex type if there was density for a core fucose and/or on the basis of mass spectroscopy assignments<sup>11,36</sup>. A core fucose was sometimes visible in one structure but not the other. We therefore interpreted glycans at some individual PNGSs as having different compositions in the two structures (**Supplementary Figs. 5–8**); such heterogeneity

is consistent with the multiple glycoforms at single PNGSs identified in preparations of BG505 SOSIP.664 protein<sup>11</sup>.

In general, the composite annealed omit electron density maps showed more density near PNGSs than the maps calculated with model phases (**Fig. 2a** and **Supplementary Figs. 5–8**); some of the extra density was not interpretable, and therefore we did not build glycan residues into these regions. In addition, some glycans were only partially ordered. Thus, a complex-type glycan could appear in our electron density maps to be a small high-mannose glycan if the core fucose and residues beyond mannose rings were disordered. In addition, the assignment of a glycan as complex type or high mannose was also not always clear because the core pentasaccharide portion of an N-glycan is common to both high-mannose and complex-type N-glycans. In other cases, our glycan interpretation was partially based on experimental data: for example, for the N301<sub>gp120</sub> glycan, a core fucose was not ordered in our electron density maps, but we modeled the N301<sub>gp120</sub> glycan as complex type in our structures, on the basis of mass spectrometry data<sup>36</sup>.

Although we built and refined glycan structures by using PDB CArbohydrate REsidue check (pdbscare) (<http://www.glycosciences.de/tools/pdb-care/>), CArbohydrate Ramachandran Plot (carp) (<http://www.glycosciences.de/tools/carp/>), and Privateer<sup>37</sup>, coordinates for glycans in our structures should be regarded as approximate (especially for regions outside the core pentasaccharide common to both complex-type and high-mannose N-glycans). The glycan density was sometimes ambiguous, as was expected given the large degree of glycan heterogeneity in HIV-1 Env, one of the most heavily glycosylated proteins known<sup>10</sup>. However, in agreement with the assumption that uninterpretable glycan density resulted from heterogeneity in the sample rather than from problems in refinement, we found unambiguous density at positions assumed to be homogenous, such as N332<sub>gp120</sub>, a high-mannose-only site<sup>11</sup>, whereas sites predicted to be more heterogeneous, such as N276<sub>gp120</sub> (refs. 11,36) (interpreted as complex type in our structures), exhibited some unassigned and uninterpretable heterogeneous electron density (**Supplementary Figs. 5–8**). Although the relatively low resolution of our crystal structures and heterogeneous glycosylation compounded inherent difficulties in making unambiguous glycan assignments, we built coordinates into very extensive densities (for example, glycan attached to N156<sub>gp120</sub>), even if the exact structure of the glycan was uncertain, in order to allow the overall extent of glycosylation at each site to be appreciated. Additional confidence in electron density interpretation came from comparing the independently refined 3.9-Å and 3.5-Å IOMA-10-1074-BG505 structures (**Table 1**). Despite caveats regarding the N-glycan coordinates, the crystal structures revealed a relatively high-resolution view of a native glycan-shield structure that could be used for assessments of the roles of complex-type N-glycans in antibody recognition and HIV-1 Env function.

By comparison with the 4.2-Å cryo-EM structure of a natively/fully glycosylated Env<sup>8</sup>, we observed a more extensively glycosylated trimer, particularly adjacent to the V3 loop and CD4bs, and the apex and gp120-gp41-interface regions of the BG505 trimer in our structures showed clusters of complex-type N-glycans not observed in fully (but not natively) glycosylated Env crystal structures<sup>9</sup> (**Fig. 2b**). We generated electron density maps at individual PNGSs and compared glycan assignments for ordered BG505 SOSIP.664 N-glycans in the 3.9-Å and 3.5-Å IOMA-10-1074-BG505 structures with glycans at analogous PNGSs in the fully and natively glycosylated 4.2-Å EM structure of a native (i.e., non-SOSIP) version of the JR-FL Env trimer<sup>8</sup>; the fully but not natively glycosylated 3.7-Å crystal structure of BG505 SOSIP.664 (i.e., high-mannose-only glycans that were not



**Figure 2** Glycan analysis. **(a)** Top, BG505 portion of IOMA-10-1074-BG505 structure with complex-type (magenta) and high-mannose-type (cyan) N-glycans shown as spheres on the 3.9-Å- and 3.5-Å-resolution structures. Middle and bottom, BG505 with  $2F_0 - F_c$  electron density contoured at  $0.8\sigma$  for ordered glycans from model-phased (middle) or composite annealed omit (bottom) maps for the 3.9-Å- and 3.5-Å-resolution structures. **(b)** Comparison of glycosylation in Env structures (side view). PGT121<sub>INT</sub> refers to the 3H + 3L intermediate that arose during maturation of PGT121 (ref. 6). Ordered N-glycans are shown as magenta (complex type) and blue (high mannose) spheres. **(c)** First and third panels, surface area accessible to a 1.4-Å probe (red) shown on natively glycosylated BG505 (3.9-Å-resolution structure; complex-type and high-mannose glycans are cyan spheres). Regions of gp120 and gp41 that are not surface accessible are shown in wheat and pink, respectively. Binding sites for IOMA (blue) and 10-1074 (green) are highlighted. Second and fourth panels, glycans displayed on the BG505 surface, with sequence conservation among 116 HIV-1 strains color-coded from white (low sequence identity) to purple (high sequence identity). Arrow points to N241<sub>gp120</sub> (yellow sphere in first panel), a PNGS in 97% of HIV-1 strains but not in BG505. The dotted ovals in the first two panels indicate an N241<sub>gp120</sub>-adjacent region of low sequence conservation that also lacks glycan density (same region shown on two adjacent protomers in each panel), which may represent an antibody-vulnerable glycan hole that would be targeted by strain-specific antibodies.

enzymatically trimmed<sup>9</sup>, a structure chosen for comparison instead of the 3.4-Å X1193.c1 fully high-mannose-glycosylated SOSIP.665 crystal structure reported in ref. 9, to compare Envs from the same HIV-1 strain); and the glycan assignments identified by mass spectrometry for natively glycosylated BG505 SOSIP.664 (ref. 11) (**Supplementary Figs. 5–8**). Although the BG505 SOSIP.664 construct<sup>34</sup> was used both for our structures and for the mass spectrometry analysis<sup>11</sup>, some differences in glycans at individual PNGSs between the mass spectrometry study and our structures might be expected if the crystals preferentially incorporated a subset of the glycosylation states within the BG505 protein.

### 10-1074 interactions with natively glycosylated Env trimer

Although N332<sub>gp120</sub> glycan-targeting bNAbs can adopt different Env-binding orientations<sup>3,38</sup>, the 10-1074 orientation resembled those of PGT121-124 family bNAbs, and, as seen in structures of BG505 bound to other PGT121-related bNAbs<sup>3,6,39</sup>, the N332<sub>gp120</sub> glycan was packed into a groove formed by 10-1074's CDRH3, CDRL1, and CDRL2 (**Fig. 3a–c**). In addition to the N332<sub>gp120</sub> glycan (Man<sub>9</sub>GlcNAc<sub>2</sub> in our 3.5-Å and 3.9-Å structures), PGT121 family bNAbs<sup>6</sup> are surrounded by glycans attached to N156<sub>gp120</sub> (modeled as complex-type tetra-antennary at 3.9-Å resolution; **Supplementary Fig. 5**), N301<sub>gp120</sub> (modeled as complex-type biantennary at 3.9-Å resolution; **Supplementary Fig. 7**), and N137<sub>gp120</sub> (disordered); thus, our structures provide what is, to our knowledge, the first visualization of the V3 loop/N332<sub>gp120</sub> glycan-dependent epitope in the context of complex-type N-glycosylation. The presence of complex-type N-glycans potentially carrying negatively charged terminal sialic acids in the vicinity of the 10-1074 epitope on Env trimers may account for the increased electropositivity of the 10-1074 combining site during affinity maturation<sup>40</sup>.

In agreement with studies demonstrating that 10-1074-lineage bNAbs are more sensitive to removal of the N332<sub>gp120</sub> glycan than PGT121-lineage bNAbs in neutralization assays<sup>30</sup>, the primary interaction for 10-1074 was with the N332<sub>gp120</sub> glycan (1,440 Å<sup>2</sup> total buried surface area (BSA)), and there were secondary or minimal interactions with the N301<sub>gp120</sub> (209 Å<sup>2</sup> BSA) and N156<sub>gp120</sub> (<40 Å<sup>2</sup> BSA) glycans, respectively (**Fig. 3c–f** and **Table 2**), and contacts as previously described with the 'GDIR' peptide motif in the gp120 V3 loop<sup>3,6,39</sup> (**Fig. 3g**). Rotation of the LC's third-framework region (FWRL3) in bound versus unbound 10-1074 (ref. 30) probably illustrates accommodation of the N301<sub>gp120</sub> glycan (**Fig. 3d,f**). Analysis of the 10-1074 interaction interface rationalizes fine specificity differences with other PGT121-family members: in PGT122, compared with 10-1074, rotation of PGT122's CDRL3 combined with an S95R<sub>LC</sub> substitution may enhance its interaction with the N156<sub>gp120</sub> glycan

(**Fig. 3e**). Indeed, PGT122 was approximately two-fold more potent against strains including the N156<sub>gp120</sub> PNGS, whereas 10-1074 showed approximately four-fold-greater potency against viral strains lacking the N156<sub>gp120</sub> PNGS (**Supplementary Fig. 9a,b**).

### IOMA interactions with natively glycosylated Env trimer

Although the Env binding orientation of IOMA differed somewhat from the orientations of both VH1-2/VRC01-class and VH1-46/8ANC131-class bNAbs (**Supplementary Fig. 10a**), the IOMA-BG505 interaction (**Fig. 4**) shared CD4-mimetic features of both VH1-2/VRC01-class and VH1-46/8ANC131-class bNAbs, including the R71<sub>HC</sub>-D368<sub>gp120</sub> interaction and CDRH2 mimicking the C'' strand of CD4 (refs. 18,19) (**Supplementary Fig. 10b,c**).

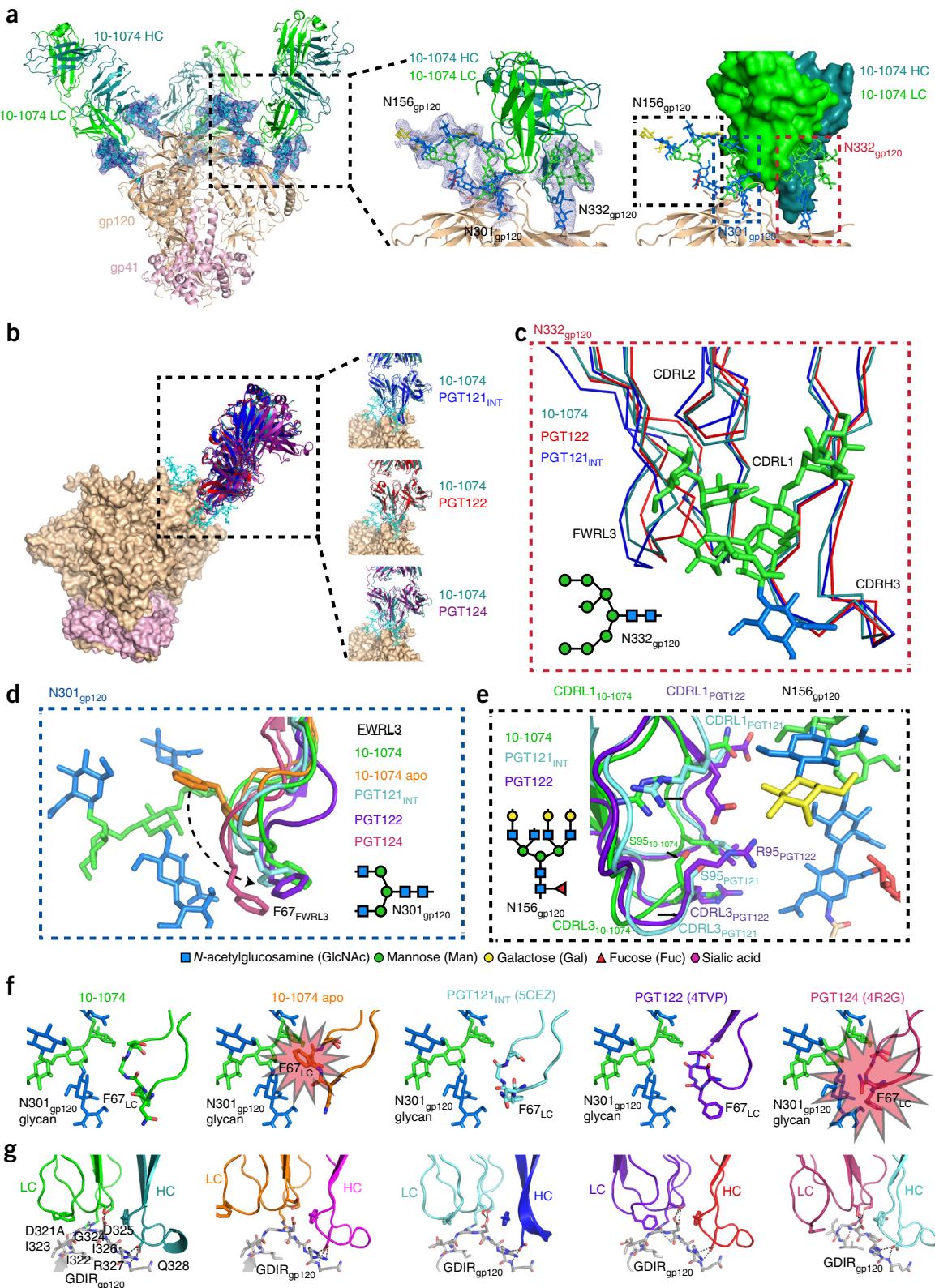
VH1-2/VRC01-class bNAbs are distinguished from VH1-46/8ANC131-like bNAbs by their five-residue CDRL3s, W50<sub>HC</sub> (VH1-2 germline encoded), and CDRH3-encoded residue W100B<sub>HC</sub><sup>29</sup>. IOMA contains W50<sub>HC</sub> and W47<sub>HC</sub>, but, in contrast to VH1-2/VRC01-class bNAbs, these residues do not contact gp120 (**Figs. 4b** and **5a**). IOMA also includes a counterpart of the signature CDRH3 W100B<sub>HC</sub> residue in VRC01-class bNAbs. In VH1-2/VRC01-class bNAbs, this tryptophan (Kabat numbering 100B; usually located four residues before the CDRH3 C terminus) is the most conserved antigen-facing residue within CDRH3, and its side chain indole nitrogen hydrogen-bonds with a side chain oxygen of N279<sub>gp120</sub> (**Fig. 4b**). The IOMA HC includes a tryptophan five residues from the C-terminal end of CDRH3 (Kabat numbering 100F) (**Supplementary Note** and **Supplementary Fig. 3a**). IOMA's W100F<sub>HC</sub> preserved the interactions observed for VRC01-class W100B<sub>HC</sub> with N279<sub>gp120</sub>/N280<sub>gp120</sub> (**Fig. 4b**).

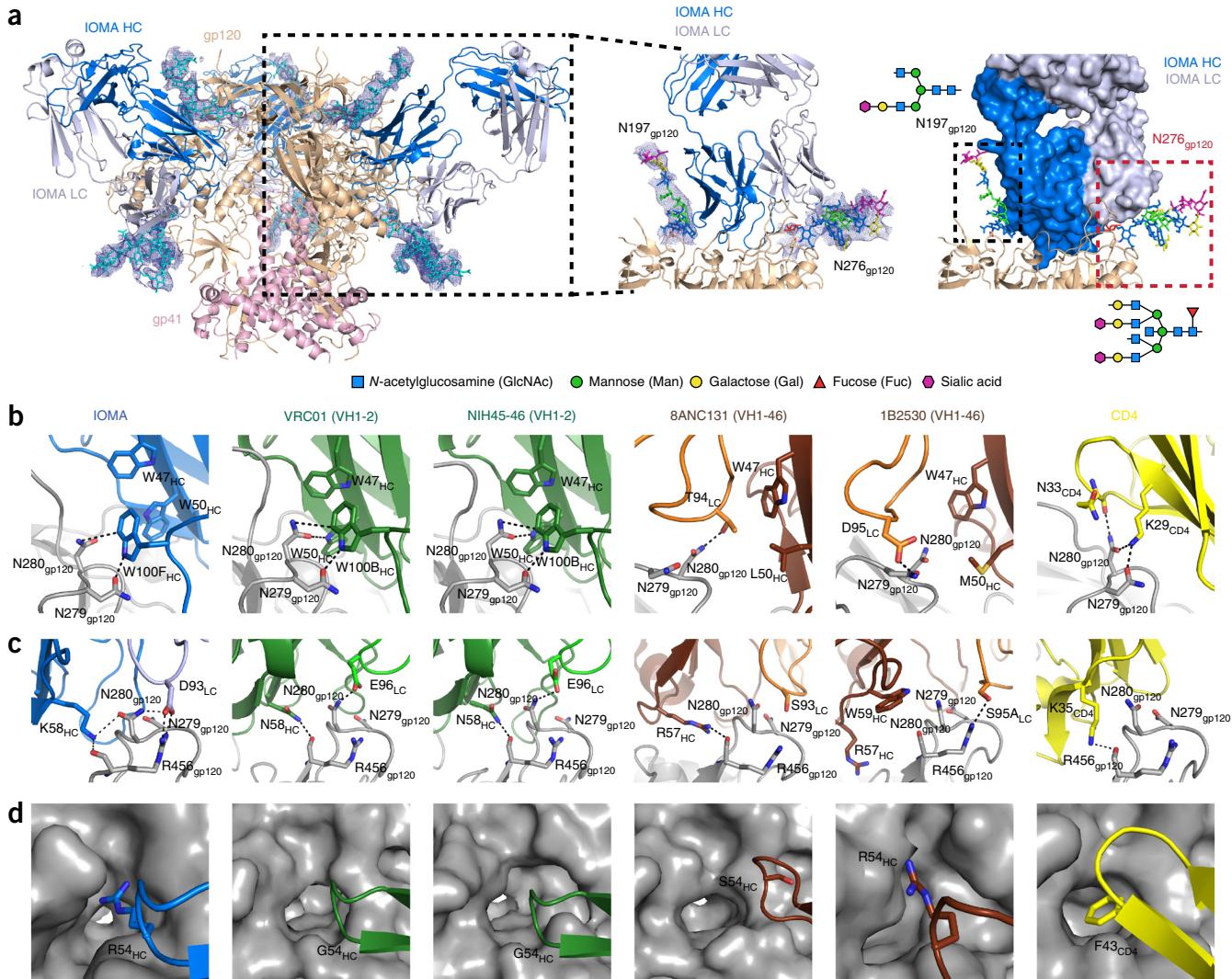
8ANC131-like bNAbs use their normal-length CDRL3s in place of a W100B<sub>HC</sub> residue to interact with N279<sub>gp120</sub>/N280<sub>gp120</sub> (**Figs. 4b** and **5b**). IOMA is unique in having both an N279<sub>gp120</sub>/N280<sub>gp120</sub>-W100B<sub>HC</sub>' interaction and a normal-length CDRL3, a combination made possible because its CDRL3 is displaced from gp120 loop D and toward the V5 loop (**Fig. 5b**). We further addressed the question of how IOMA recognizes Env with a normal-length CDRL3 together with signature VRC01-class residues by comparing Env interactions with IOMA, VH1-2/VRC01-class, and VH1-46/8ANC131-class bNAbs. In 8ANC131-class bNAbs, the longer CDRL3 interacts with N279<sub>gp120</sub>/D279<sub>gp120</sub> (**Fig. 4b**), and the C-terminal portion of CDRH3 is farther from loop D (**Fig. 5b**). In VRC01-class bNAbs, E96<sub>LC</sub> (within a five-residue CDRL3) hydrogen-bonds with N280<sub>gp120</sub> (**Fig. 4c**) and makes a backbone contact to G459<sub>gp120</sub> (**Fig. 5c**). IOMA's eight-residue CDRL3 also includes a negatively charged residue, D93<sub>HC</sub>, that interacts with R456<sub>gp120</sub> and N280<sub>gp120</sub> (**Figs. 4c** and **5c**). However, IOMA's longer CDRL3 is accommodated by a shift in BG505's gp120 V5 loop, relative to its position in gp120s complexed with VH1-2/VRC01-class bNAbs,

**Figure 3** 10-1074 interactions with BG505. Glycans shown are from the IOMA-10-1074-BG505 structure and are colored as indicated. **(a)** 10-1074 Fab (green) bound to BG505 (gp120 in wheat; gp41 in pink). Glycans near the binding site are shown in stick representation with dark-blue 2F<sub>0</sub> – F<sub>c</sub> electron density contoured at 0.8σ (left and middle) or as sticks alone (right). Middle and right panels show close-up views of the boxed area in the left panel, with the Fab domains as ribbons (middle) or in space-filling representation (right). **(b)** Comparison of binding orientations of 10-1074 and other PGT121 family members. PGT121<sub>INT</sub> refers to the 3H + 3L intermediate that arose during maturation of PGT121 (ref. 6). Orientation for PGT124 was generated by modeling the gp120 of a PGT124-gp120 structure (PDB 4R2G) onto BG505. **(c–e)** Structural overlays of 10-1074 and other PGT121 family members in regions near the glycans attached to N332<sub>gp120</sub> (**c**), N301<sub>gp120</sub> (**d**), and N156<sub>gp120</sub> (**e**). 10-1074 apo is the structure of the unbound Fab (PDB 4FQ2). **(f)** Observed and potential interactions with N301<sub>gp120</sub> glycan. F67<sub>LC</sub> in the LC of the BG505-bound 10-1074 Fab is shifted from its position in the unbound 10-1074 Fab structure (10-1074 apo; PDB 4FQ2) to accommodate the N301<sub>gp120</sub> glycan. The red starburst in the model of 10-1074 apo aligned with 10-1074 from the IOMA-10-1074-BG505 structure indicates a predicted clash. F67<sub>LC</sub> from the LCs of the PGT121<sub>INT</sub> and PGT122 Fabs in complex structures with BG505 could accommodate the N301<sub>gp120</sub> glycan if the glycan were oriented as in the IOMA-10-1074-BG505 structure, but a clash (red starburst) is predicted between F67<sub>LC</sub> from the LC of PGT124 and the N301<sub>gp120</sub> glycan unless the glycan is shifted from this position. **(g)** Interactions with GDIR motif. Similarly to other PGT121-124 family bNAbs<sup>3,6,39</sup>, 10-1074 reaches through the glycan shield and contacts the GDIR peptide motif (G324<sub>gp120</sub>-D325<sub>gp120</sub>-I326<sub>gp120</sub>-R327<sub>gp120</sub>) in the gp120 V3 loop. PGT121<sub>INT</sub> refers to the 3H + 3L intermediate that arose during maturation of PGT121 (ref. 6).

that allows IOMA's CDRL3 to penetrate the groove between the V5 and D loops (Figs. 4c and 5d). IOMA's D93<sub>LC</sub>-R456<sub>gp120</sub> interaction is not found in VRC01-class bNAbs, wherein N58<sub>HC</sub> interacts with the backbone of R456<sub>gp120</sub>. In IOMA, the VH1-2 germline residue N58<sub>HC</sub> is mutated to K58<sub>HC</sub>, which interacts with the backbone

of N280<sub>gp120</sub>, a similar interaction to that in 8ANC131 (Fig. 4c). The hydrophobic gp120 pocket, which normally accommodates F43<sub>CD4</sub> (ref. 2) and has been targeted by engineered substitutions of G54<sub>HC</sub> in VRC01-class bNAbs<sup>14</sup>, is filled by R54<sub>HC</sub> (Fig. 4d), as seen in a gp120-complex structure with 1B2530, a VH1-46-derived bNAb<sup>19</sup>.





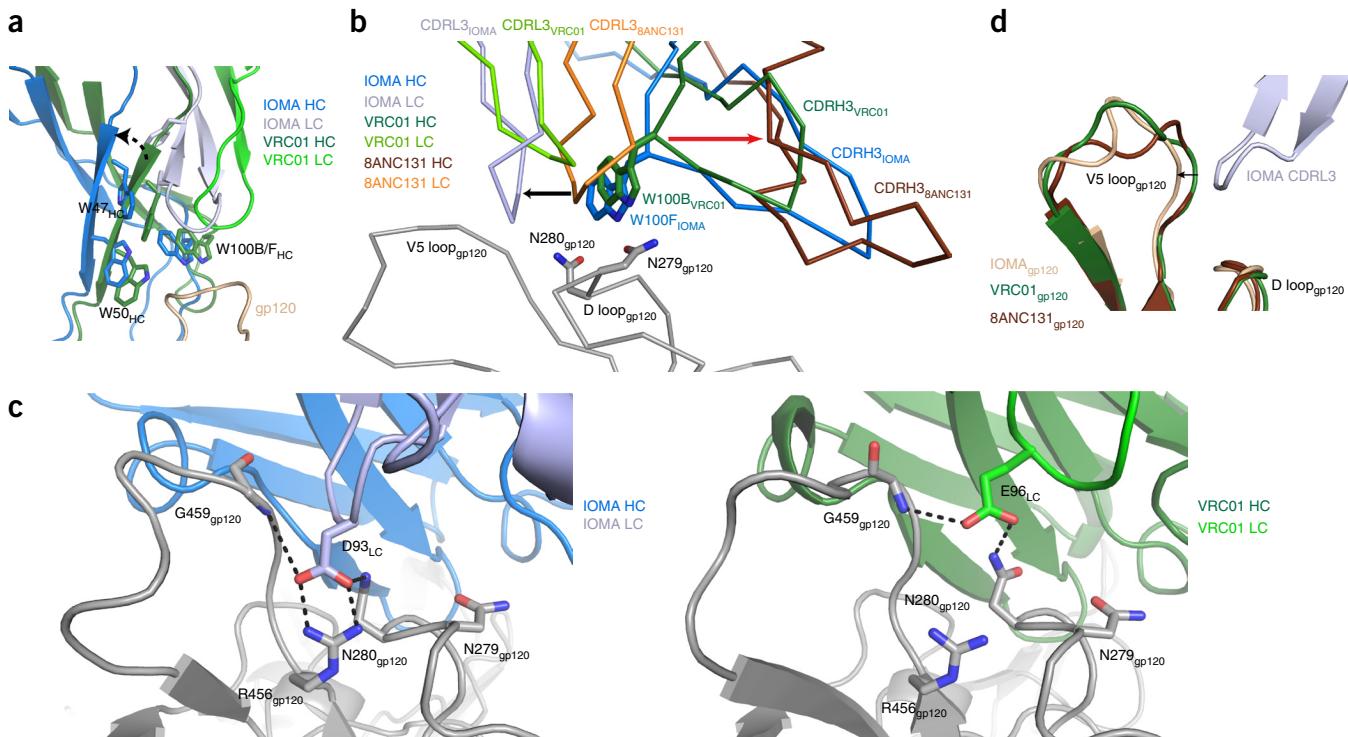
**Figure 4** IOMA interactions with BG505. **(a)** IOMA Fabs (blue) bound to BG505 (gp120 in wheat; gp41 in pink). Glycans near the binding site are shown as sticks with dark-blue  $2F_0 - F_c$  electron density contoured at  $0.8\sigma$  (left and center) or as sticks alone (right; color-coded as in schematics and in Fig. 3). **(b)** Interactions of W100B<sub>HC</sub> (W100F<sub>HC</sub> in IOMA) with N279<sub>gp120</sub>/N280<sub>gp120</sub> in indicated bNAbs and CD4. **(c)** Interactions of CDRL3s with N279<sub>gp120</sub>/N280<sub>gp120</sub>. **(d)** Interactions with the F43<sub>CD4</sub> pocket on gp120.

The IOMA Fab is framed on both sides by Env N197<sub>gp120</sub> and N276<sub>gp120</sub> glycans, interpreted as complex-type biantennary and tetra-antennary, respectively, in the 3.9-Å structure, with

minor interactions with a high-mannose glycan at N363<sub>gp120</sub> (Supplementary Fig. 10d and Table 2). IOMA accommodates the N276<sub>gp120</sub> glycan with a short  $\alpha$ -helix in CDRL1 instead of the extended loop in other CD4-mimetic bNAbs (Fig. 6a–d). The need for CDRL1 to accommodate the glycan attached to N276<sub>gp120</sub> is a barrier to the development of VRC01-class bNAbs<sup>19,20</sup> that is typically overcome by CDRL1 deletions or by SHM introducing multiple glycine residues<sup>20</sup> (Fig. 6e). The VL2-23-derived CDRL1 of IOMA has no indels and is two or three residues longer than the CDRL1 of the common VRC01-class LCs KV1-33 and KV3-20, and IOMA's CDRL1 acquired only one additional glycine residue during SHM (G29<sub>LC</sub>). The short  $\alpha$ -helix in IOMA's CDRL1 is a CDRL1 conformation not observed in available structures of VRC01-class or 8ANC131-class bNAbs (Fig. 6a–d). However, this helical conformation (designated L1-14-02 in refs. 41,42) is common in the CDRL1s of antibodies derived from human germline VL2-23 and the closely related VL2-14 and VL2-8 germlines. For example, gp41-targeting bNAbs 3BC176 and 3BC315, which share germline VH and VL genes with IOMA, have CDRL1 loops in the same helical conformation as

**Table 2** Buried surface areas at the indicated interfaces

Interface	Total	Buried surface area (BSA) ( $\text{\AA}^2$ )		
		Ab	BG505	Glycan
BG505-10-1074	2,697	1,219	508	970
HC	1,308	597	210	501
LC	1,389	622	298	469
BG505-IOMA	3,554	1,787	1,003	764
HC	2,435	1,230	815	390
LC	1,119	557	188	374
Without N363 <sub>gp120</sub> glycan				
BG505-IOMA	3,396	1,706	1,003	687
HC	2,777	1,149	815	313
LC	1,119	557	188	374



**Figure 5** Comparison of CDRL3 in IOMA and other CD4-mimetic bNAb structures. **(a)** Shift in IOMA's FWRH2  $\beta$ -strand. To accommodate IOMA's normal-length CDRL3, there is a shift relative to VRC01 in IOMA's FWRH2  $\beta$ -strand; thus, IOMA residues W47<sub>HC</sub> and W50<sub>HC</sub> do not interact with gp120. **(b)** CDRL3 comparisons in CD4-mimetic antibody–Env complexes (aligned on the gp120s from the IOMA–10-1074–BG505 (shown in gray) and the VRC01–gp120 (PDB 3NGB) and 8ANC131–gp120 (PDB 4RWY) structures). CDRL3s are light blue (IOMA), bright green (VRC01), and orange (8ANC131). The black arrow shows displacement of the eight-residue IOMA CDRL3 relative to the five-residue VRC01 CDRL3 away from W100F<sub>HC-IOMA</sub> and toward the gp120 V5 loop. The red arrow shows displacement of the CDRH3 in VH1-46-derived 8ANC131-like bNAbs (brown) relative to VRC01 class (dark green) and IOMA (dark blue) bNAbs. **(c)** Comparison of structurally analogous acidic residues within the CDRL3s of IOMA (D93<sub>LC</sub>) and VRC01 (E96<sub>LC</sub>). Left, interface between gp120 (gray) and IOMA (HC, dark blue; LC, light blue), demonstrating interactions between IOMA D93<sub>LC</sub> and gp120 residues R456<sub>gp120</sub>, G459<sub>gp120</sub>, and N280<sub>gp120</sub>. Right, interface between gp120 (gray) and VRC01 (HC, dark green; LC, light green) demonstrating interactions between VRC01 E96<sub>LC</sub> and residues G459<sub>gp120</sub> and N280<sub>gp120</sub>. **(d)** V5-loop shift. To accommodate IOMA's normal-length CDRL3, there is a shift in BG505's V5 loop in the IOMA–10-1074–BG505 structure compared with V5 loops in complex structures of gp120s bound to VRC01 and 8ANC131.

in IOMA (PDB 5AWN and 5CCK)<sup>43</sup>. VL2-14-derived bNAbs PG9, PG16, and 35O22 (PDB 3U4E, 4DQO, and 4TVP)<sup>4,44,45</sup> also include  $\alpha$ -helical CDRL1s. Although VRC01-class bNAb VRC-PG20 derives from VL2-14, its CDRL1 has a six-residue deletion and does not have a helical conformation<sup>20</sup>.

In contrast with observed interactions in structures of VH1-2/VRC01-class bNAb–Env complexes<sup>13,14,18–20</sup>, IOMA interacts extensively with the Env N276<sub>gp120</sub> and N197<sub>gp120</sub> glycans. Of the 1,839 Å<sup>2</sup> BSA, 749 Å<sup>2</sup> involved these glycans (Table 2), which are either not present in gp120 constructs used for crystallization (N197<sub>gp120</sub>) or disordered (N276<sub>gp120</sub>) in most VRC01-class–gp120 structures. To compare how IOMA and other VH1-2-derived bNAbs accommodate these glycans, we evaluated the neutralization potencies of IOMA and VRC01-class bNAbs against HIV-1<sup>YU2</sup> pseudoviruses with N197<sub>gp120</sub> or N276<sub>gp120</sub> glycan deletions (Supplementary Fig. 9c,d). The results suggested that glycosylation at N197<sub>gp120</sub> (98% conserved) is accommodated better by IOMA than by the VRC01-class bNAbs VRC01 or NIH45-46, with both IOMA and the VRC01-class bNAbs showing a slight preference for removal of the N276<sub>gp120</sub> glycan (95% conserved). IOMA's similar ability to that of VRC01 for accommodating the N276<sub>gp120</sub> glycan, one of the primary roadblocks in the development of VRC01-class bNAbs<sup>20,32</sup>, required relatively small changes from its germline CDRL1 sequence (Fig. 6e). This aspect may favor the development of IOMA-like bNAbs and relate to IOMA's distinct Env-binding orientation (Supplementary Fig. 10a) and  $\alpha$ -helical CDRL1 (Fig. 6a).

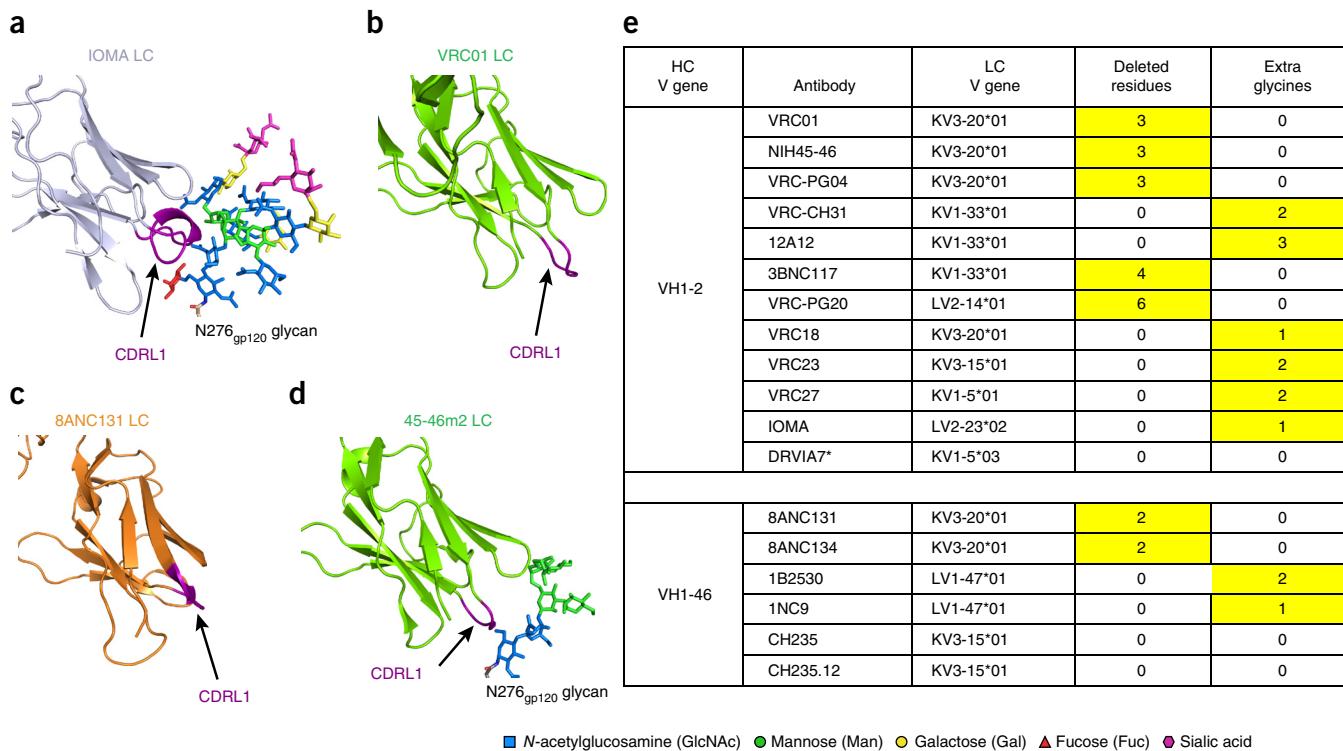
## DISCUSSION

### Gaps in the glycan shield

The 10-1074 and IOMA footprints are apparent as contiguous solvent-accessible surface areas when mapped onto the natively glycosylated BG505 structure (Fig. 2c, green and blue). The footprints fall in Env regions that are relatively conserved in sequence (Fig. 2c, dark purple), thus rationalizing why both antibodies are broadly neutralizing. These results are consistent with those from a recent study reporting that neutralizing antibodies target glycan-deficient patches; i.e., holes in the glycan shield<sup>46</sup>. Potential antibody epitopes that elicit strain-specific, but not broadly neutralizing, antibodies should be apparent as holes in the glycan shield (Fig. 2c, red) in which the exposed surface exhibits low sequence conservation (Fig. 2c, white or light purple). A glycan-deficient region adjacent to N241<sub>gp120</sub> (a PNGS in 97% of HIV-1 strains, but not BG505) (Fig. 2c, dotted oval) also shows low sequence conservation, thus suggesting that this region represents a hole in the glycan shield that could be targeted by autologous but not broadly neutralizing antibodies.

### Implications for vaccine design to elicit CD4bs bNAbs

VH1-2/VRC01-class bNAbs are promising targets for vaccine design because they evolved in multiple donors and recognize HIV-1 Env through conserved interactions<sup>14,18–20,28,29,33</sup>. The IOMA-BG505 structure demonstrating that VH1-2-derived CD4-mimetic bNAbs



**Figure 6** Comparison of CDRL1 in IOMA and other CD4-mimetic bNAb structures. (a) LC of IOMA showing  $\alpha$ -helical CDRL1 and N276<sub>gp120</sub> glycan from the IOMA-10-1074-BG505 structure. (b) LC of VRC01 from a VRC01-gp120 structure (PDB 3NGB). The N276<sub>gp120</sub> glycan was disordered in this structure. (c) LC of 8ANC131 from an 8ANC131-gp120 structure (PDB 4RWY). The N276<sub>gp120</sub> glycan was disordered in this structure. (d) LC from 45-46m2 (an engineered VRC01-class bNAb<sup>13</sup>) from a 45-46m2-gp120 structure (PDB 4JKP). The N276<sub>gp120</sub> glycan was partially ordered in this structure. Glycans are shown as sticks; color-coded as shown in key. (e) Changes in CDRL1 versus germline for CD4-mimetic bNAbs. The mechanism by which each CDRL1 accommodates the N276<sub>gp120</sub> glycan is highlighted in yellow. The reduced breadth and potency of DRVIA7 compared with other VRC01-class bNAbs appears to be due to its LC<sup>32</sup>.

are not limited to five-residue CDRL3s suggests an additional pathway for vaccine elicitation by immunogens targeting VRC01-class bNAbs. To elicit VRC01-class bNAbs, immunogens have been designed to bind to inferred VH1-2 germline precursors<sup>21–25</sup>. In transgenic mice expressing germline VRC01-class HCs, the designed immunogens select VH1-2 B-cells with short CDRL3s<sup>22,27</sup>, although relatively few clones with five-residue CDRL3s have been isolated, thus reflecting their low frequency (~1% of LCs)<sup>29</sup>. The existence of bNAbs that combine features of VRC01-like and 8ANC131-like classes has implications for immunization strategies targeting VRC01-like bNAbs. First, VRC01-class-targeting immunogens may also initially elicit IOMA-like antibodies; thus, studies using these immunogens should examine whether IOMA-like antibodies (i.e., VH1-2 Abs with longer CDRL3s) are also elicited. Indeed, isolation of naive B cells binding to the VRC01-class germline-targeting immunogen eOD-GT8 was found to enrich VH1-2-expressing cells from 4% to 50%, but 70% of eOD-GT8-binding/VH1-2-expressing cells do not express an antibody with a five-residue CDRL3 (ref. 25); some or even most of these may represent B cells with IOMA-like antibody receptors.

In the context of eliciting CD4-mimetic bNAbs, IOMA represents an alternative pathway from VRC01-class bNAbs. Although this pathway may not lead to bNAbs with the breadth and potency of VRC01-like bNAbs, it is possible that the IOMA pathway may more readily lead to an effective vaccine response, owing to higher frequencies of normal-length CDRL3s compared with the rare five-residue CDRL3s required for VRC01-class bNAbs<sup>29</sup>, and a lower need for SHM. The isolation of fewer IOMA-like bNAbs than VRC01-class bNAbs to date

may have resulted from only the most broadly neutralizing sera having been investigated. Regardless of whether elicitation of IOMA-like antibodies is desirable for VH1-2-targeting immunogens, knowledge of this binding mode opens the possibility of immunogen design to either favor or disfavor this pathway.

## METHODS

Methods and any associated references are available in the online version of the paper.

**Accession codes.** Coordinates and structure factors for BG505 SOSIP.664 in complex with IOMA and 10-1074 Fab have been deposited in the Protein Data Bank under accession codes PDB 5T3Z (3.5-Å resolution) and PDB 5T3X (3.9-Å resolution). IOMA HC and LC gene sequences have been deposited in GenBank under accession codes KX610770 and KX610771, respectively).

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

## ACKNOWLEDGMENTS

We thank L. Stamatatos, M. Crispin, and A.-J. Behrens for helpful discussions; J. Vielmetter, J. Nhan, and the Caltech Protein Expression Center for producing proteins; P. Gnanapragasam for performing in-house neutralization assays; J. Kaiser for assistance with data processing; K. Vellizon and N. Thomas for assistance with FACS sorting; T. Oliveira for bioinformatic analyses; M. Murphy and H. Wang for assistance in making figures; the beamline staff at SSRL; A. Hurley (Rockefeller University Hospital) and C. Lehmann and G. Kremer (Infectious Disease Division at the University Hospital Cologne) for supporting the collection of patient material; and members of the Bjorkman and Nussenzweig laboratories

for critical reading of the manuscript. This research was supported by National Institute Of Allergy And Infectious Diseases of the National Institutes of Health grant HIVRAD P01 AI100148 (P.J.B.) (the content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health), the Bill and Melinda Gates Foundation (Collaboration for AIDS Vaccine Discovery Grant OPP1124068 (M.C.N. and P.J.B.)), Comprehensive Antibody-Vaccine Immune Monitoring Consortium grant 1032144 (M.S.S.), the Rockefeller University Center for Clinical and Translational Science grant UL1 TR000043/KL2TR000151 from the National Center for Advancing Translational Sciences (L.v.B.), and the Molecular Observatory at Caltech, supported by the Gordon and Betty Moore Foundation. F.K. is supported by the Heisenberg-Program of the DFG (KL 2389/2-1), the European Research Council (ERC-StG639961), and the German Center for Infection Research (DZIF), partner site Bonn-Cologne, Cologne, Germany. Operations at the Stanford Synchrotron Radiation Lightsource are supported by the US Department of Energy and the National Institutes of Health.

#### AUTHOR CONTRIBUTIONS

H.B.G., L.v.B., A.P.W. Jr., M.C.N., and P.J.B. conceived the experiments; H.B.G. solved and analyzed crystal structures; L.v.B. performed single-cell experiments and isolated, cloned, and analyzed R1 antibodies; A.P.W. Jr. performed computational and bioinformatics analyses of VRC01-class, IOMA-class, and 8ANC131-class antibodies; H.B.G. and M.S. purified proteins for crystallography and optimized crystals; A.G. and J.G. produced monoclonal antibodies; M.S.S. conducted *in vitro* neutralization assays; F.K. and G.F. were the principal investigators for the clinical protocols and provided serum and PBMCs of subject R1; H.B.G., L.v.B., A.P.W., M.C.N. and P.J.B. wrote the paper, on which all principal investigators commented.

#### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Lasky, L.A. *et al.* Neutralization of the AIDS retrovirus by antibodies to a recombinant envelope glycoprotein. *Science* **233**, 209–212 (1986).
- Kwong, P.D. *et al.* Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648–659 (1998).
- Julien, J.P. *et al.* Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1477–1483 (2013).
- Pancera, M. *et al.* Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* **514**, 455–461 (2014).
- Scharf, L. *et al.* Broadly neutralizing antibody 8ANC195 recognizes closed and open states of HIV-1 Env. *Cell* **162**, 1379–1390 (2015).
- Garcés, F. *et al.* Affinity maturation of a potent family of HIV antibodies is primarily focused on accommodating or avoiding glycans. *Immunity* **43**, 1053–1063 (2015).
- Lee, J.H., de Val, N., Lyumkis, D. & Ward, A.B. Model building and refinement of a natively glycosylated HIV-1 Env protein by high-resolution cryoelectron microscopy. *Structure* **23**, 1943–1951 (2015).
- Lee, J.H., Ozorowski, G. & Ward, A.B. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. *Science* **351**, 1043–1048 (2016).
- Stewart-Jones, G.B.E. *et al.* Trimeric HIV-1-Env structures define glycan shields from clades A, B, and G. *Cell* **165**, 813–826 (2016).
- Doores, K.J. The HIV glycan shield as a target for broadly neutralizing antibodies. *FEBS J.* **282**, 4679–4691 (2015).
- Behrens, A.J. *et al.* Composition and antigenic effects of individual glycan sites of a trimeric HIV-1 envelope glycoprotein. *Cell Rep.* **14**, 2695–2706 (2016).
- Binley, J.M. *et al.* Role of complex carbohydrates in human immunodeficiency virus type 1 infection and resistance to antibody neutralization. *J. Virol.* **84**, 5637–5655 (2010).
- Diskin, R. *et al.* Restricting HIV-1 pathways for escape using rationally designed anti-HIV-1 antibodies. *J. Exp. Med.* **210**, 1235–1249 (2013).
- Diskin, R. *et al.* Increasing the potency and breadth of an HIV antibody by using structure-based rational design. *Science* **334**, 1289–1293 (2011).
- Kwon, Y.D. *et al.* Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat. Struct. Mol. Biol.* **22**, 522–531 (2015).
- Kong, L. *et al.* Complete epitopes for vaccine design derived from a crystal structure of the broadly neutralizing antibodies PGT128 and 8ANC195 in complex with an HIV-1 Env trimer. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 2099–2108 (2015).
- Scharf, L. *et al.* Antibody 8ANC195 reveals a site of broad vulnerability on the HIV-1 envelope spike. *Cell Rep.* **7**, 785–795 (2014).
- Zhou, T. *et al.* Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329**, 811–817 (2010).
- Zhou, T. *et al.* Structural repertoire of HIV-1-neutralizing antibodies targeting the CD4 supersite in 14 donors. *Cell* **161**, 1280–1292 (2015).
- Zhou, T. *et al.* Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity* **39**, 245–258 (2013).
- Jardine, J. *et al.* Rational HIV immunogen design to target specific germline B cell receptors. *Science* **340**, 711–716 (2013).
- Jardine, J.G. *et al.* Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen. *Science* **349**, 156–161 (2015).
- McGuire, A.T. *et al.* Antigen modification regulates competition of broad and narrow neutralizing HIV antibodies. *Science* **346**, 1380–1383 (2014).
- McGuire, A.T. *et al.* Specifically modified Env immunogens activate B-cell precursors of broadly neutralizing HIV-1 antibodies in transgenic mice. *Nat. Commun.* **7**, 10618 (2016).
- Jardine, J.G. *et al.* HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. *Science* **351**, 1458–1463 (2016).
- McGuire, A.T. *et al.* Engineering HIV envelope protein to activate germline B cell receptors of broadly neutralizing anti-CD4 binding site antibodies. *J. Exp. Med.* **210**, 655–663 (2013).
- Dosenovic, P. *et al.* Immunization for HIV-1 broadly neutralizing antibodies in human Ig knockin mice. *Cell* **161**, 1505–1515 (2015).
- Wu, X. *et al.* Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329**, 856–861 (2010).
- West, A.P. Jr., Diskin, R., Nussenzweig, M.C. & Bjorkman, P.J. Structural basis for germ-line gene usage of a potent class of antibodies targeting the CD4-binding site of HIV-1 gp120. *Proc. Natl. Acad. Sci. USA* **109**, E2083–E2090 (2012).
- Mouquet, H. *et al.* Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. *Proc. Natl. Acad. Sci. USA* **109**, E3268–E3277 (2012).
- Scheid, J.F. *et al.* Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458**, 636–640 (2009).
- Bonsignori, M. *et al.* Maturation pathway from germline to broad HIV-1 neutralizer of a CD4-mimic antibody. *Cell* **165**, 449–463 (2016).
- Scheid, J.F. *et al.* Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333**, 1633–1637 (2011).
- Sanders, R.W. *et al.* A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog.* **9**, e1003618 (2013).
- Adams, P.D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
- Go, E.P. *et al.* Characterization of glycosylation profiles of HIV-1 transmitted/founder envelopes by mass spectrometry. *J. Virol.* **85**, 8270–8284 (2011).
- Aguirre, J. *et al.* Privateer: software for the conformational validation of carbohydrate structures. *Nat. Struct. Mol. Biol.* **22**, 833–834 (2015).
- Kong, L. *et al.* Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. *Nat. Struct. Mol. Biol.* **20**, 796–803 (2013).
- Pejchal, R. *et al.* A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. *Science* **334**, 1097–1103 (2011).
- Scharf, L. *et al.* Structural basis for germline antibody recognition of HIV-1 immunogens. *eLife* **5**, e13783 (2016).
- Adolf-Bryfogle, J., Xu, Q., North, B., Lehmann, A. & Dunbrack, R.L. Jr. PylgClassify: a database of antibody CDR structural classifications. *Nucleic Acids Res.* **43**, D432–D438 (2015).
- North, B., Lehmann, A. & Dunbrack, R.L. Jr. A new clustering of antibody CDR loop conformations. *J. Mol. Biol.* **406**, 228–256 (2011).
- Lee, J.H. *et al.* Antibodies to a conformational epitope on gp41 neutralize HIV-1 by destabilizing the Env spike. *Nat. Commun.* **6**, 8167 (2015).
- McLellan, J.S. *et al.* Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480**, 336–343 (2011).
- Pancera, M. *et al.* Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. *Nat. Struct. Mol. Biol.* **20**, 804–813 (2013).
- Crooks, E.T. *et al.* Vaccine-elicited tier 2 HIV-1 neutralizing antibodies bind to quaternary epitopes involving glycan-deficient patches proximal to the CD4 binding site. *PLoS Pathog.* **11**, e1004932 (2015).

## ONLINE METHODS

**Patient samples.** Subject R1 is an HIV-1-infected individual who started antiretroviral therapy (ART) four months after initial diagnosis. ART treatment was paused twice for 2.5 years at 17 months and 7 years after treatment initiation, and for 3 months approximately 11 years after initiation. A serum sample was collected under informed written consent and in accordance with the University of Cologne Institutional Review Board (09–281). The serum sample was heat-inactivated for 1 h at 56 °C, and the IgG fraction was purified with Protein G-Sepharose 4 Fast Flow (GE Healthcare). Sterile filtration and buffer exchange to PBS was performed before testing for neutralizing activity. IgG was screened for neutralizing activity against a panel of tier 1, tier 1B, and tier 2 viruses representing eight different clades or interclade recombinants. A leukapheresis sample (17 years after initial HIV-1 diagnosis) was obtained under informed consent and under approval of the Rockefeller University Institutional Review Board (MNU-0628). The sample was processed within 2 h of collection. Serum and plasma samples were stored, and PBMCs were isolated by density gradient centrifugation. The absolute number of peripheral blood mononuclear cells was determined with an automated cell counter (Vi-Cell XR; Beckman Coulter), and cells were cryopreserved in FBS/10% DMSO.

**Single-B-cell sorting.** gp140<sup>YU2</sup> and 2cc-core were used for sorting experiments with R1 and were produced as previously described<sup>33</sup>. Purified total B cells were stained with combinations of anti-human CD19 (FITC human CD19 (BD, 340864)), IgG (APC human IgG (BD, 550931)) antibodies and PE-labeled gp140<sup>YU2</sup> or 2cc-core. Antigen-specific IgG<sup>+</sup> memory B cells were single-cell-sorted on a BD Aria cytometer into 96-well PCR plates containing 4 µL/well of lysis buffer. Plates were immediately frozen on dry ice before storage at –80 °C.

**Amplification and cloning of antibody variable genes.** For each cell, cDNA was generated by two-step reverse transcription with random primers. The sequences of the V<sub>H</sub> and V<sub>L</sub> domains were amplified by nested PCR, with a combination of previously described primers<sup>33,47,48</sup>. Sequence analysis was performed to identify Ig gene usage, HC and LC CDR3s, and the number of V<sub>H</sub>-V<sub>L</sub> somatic hypermutations (IgBLAST, <http://www.ncbi.nlm.nih.gov/igblast/> and IMGT, <http://www.imgt.org/>). For cloning by PCR, the primers contained 15–20 base pairs of 5' and 3' homology to restriction-digested expression vectors to allow for cloning by homologous recombination into DH5α, as previously described<sup>49</sup>.

**In vitro neutralization assays.** IgGs for neutralization assays were expressed in HEK293-6E cells (National Research Council of Canada) by transient transfection of equal amounts of HC and LC vectors and purified by Protein G-Sepharose 4 Fast Flow Chromatography (GE Healthcare) from transfected cell supernatants collected after 7 d of culture. Sterile filtration and buffer exchange to PBS were performed before testing for neutralizing activity.

Pseudovirus neutralization was monitored by the reduction of HIV-1 Tat-induced luciferase reporter gene expression in the presence of a single round of pseudovirus infection in TZM-bl cells, as previously described<sup>50</sup>. Neutralization assays were conducted by the Collaboration for AIDS Vaccine Discovery (CAVD) core neutralization facility (Fig. 1 and Supplementary Figs. 1 and 2) or in house (Supplementary Fig. 9). Data were fit in Prism (GraphPad) with nonlinear regression to derive IC<sub>50</sub> values. The IC<sub>50</sub> values were derived from independent replicates of manual and robotic assays conducted with eight potential inhibitor concentrations tested in duplicate or triplicate, and generally agreed within two-to four-fold. Average IC<sub>50</sub> values reported in the figures and tables are geometric means calculated with the formula  $(\prod a_i)^{(1/N)}$ ;  $i = 1, 2, \dots, N$ . Geometric means are suitable statistics for data sets covering multiple orders of magnitude<sup>51</sup>, as is the case for neutralization data across multiple viral strains. Fold improvements were calculated as the ratio of the geometric mean IC<sub>50</sub> values for the reagents being compared.

**Protein production and purification for crystallography.** Fabs from the IOMA and 10-1074 bNabs (in which the 10-1074 Fab included a C263S<sub>LC</sub> substitution<sup>52</sup>) were expressed and purified as described in previous studies<sup>5</sup>. Briefly, Fabs were expressed by transient transfection in HEK293-6E cells (obtained from the National Research Council of Canada; tested for mycoplasma) with expression vectors containing genes encoding the LC and a C-terminally histidine-tagged Fab portion of the HC. Histidine-tagged Fabs were purified with

Ni<sup>2+</sup>-NTA affinity chromatography (GE Healthcare) and SEC with a Superdex 200 16/60 column (GE Healthcare).

A gene encoding BG505 SOSIP.664, a soluble clade A gp140 trimer<sup>34</sup>, was constructed to include ‘SOS’ substitutions (A501C<sub>gp120</sub> and T605C<sub>gp41</sub>), the ‘IP’ substitution (I559P<sub>gp41</sub>), the N-linked glycan sequence at residue 332<sub>gp120</sub> (T332N<sub>gp120</sub>), an enhanced gp120-gp41 cleavage site (REKR to RRRRRR), and a stop codon after residue 664<sub>gp41</sub> (Env numbering according to HX nomenclature). Protein was expressed in HEK293-6E cells (National Research Council of Canada) by transient transfection of plasmids encoding BG505 SOSIP and soluble furin at a ratio of 4:1, as previously described<sup>5</sup>, except that the cells were not treated with kifunensine. BG505 SOSIP protein was isolated from cell supernatants with a 2G12 immunoaffinity column made by covalently coupling 2G12 IgG monomer to an NHS-activated Sepharose column (GE Healthcare). After elution with 3 M MgCl<sub>2</sub> and immediate buffer exchange into Tris-buffered saline, pH 8.0 (TBS), trimers were purified with Superdex 200 16/60 SEC (GE Healthcare). The trimer fractions were pooled and repurified with the same column. Twelve 1.0-mL fractions were collected, which were pooled in groups of two and then stored separately. Selected fractions were combined with IOMA and 10-1074 Fabs (Supplementary Fig. 4), thus resulting in two crystal structures: a 3.9-Å-resolution structure from BG505 fractions 7 and 8, and a 3.5-Å-resolution structure from BG505 fractions 11 and 12 (Table 1).

**Crystallization.** Samples for crystallization were produced by incubation of BG505 SOSIP with a 1:1:1 molar ratio of IOMA and 10-1074 Fabs for ~16 h at room temperature. The resulting complex was concentrated to 5–10 mg/ml with a 30-kDa concentrator (Amicon). Crystal trials were carried out with the sitting-drop vapor-diffusion method at room temperature by equilibration of equal volumes of the protein complex solution and reservoir solution with a TTP LabTech Mosquito robot and commercially purchased kits (Hampton Research). Crystals of IOMA–BG505–10-1074 complex (one BG505 protomer per asymmetric unit) were obtained by combining 0.2 µL of protein sample with 0.2 µL of 200 mM ammonium citrate tribasic, pH 7.0, 100 mM imidazole, pH 7.0, and 20% PEG MME 2000 at 20 °C. Crystals were cryoprotected in mother liquor supplemented with 20% glycerol and incubated for at least 10 min before being flash frozen in liquid nitrogen.

**Crystallographic data collection, structure determination, and refinement.** All structures were solved with data sets collected at 100 K and 1-Å resolution on Beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL) with a Pilatus 6M pixel detector (Dectris). Data sets were indexed, integrated with MOSFLM<sup>53</sup>, and then merged with AIMLESS in the CCP4 software package<sup>54</sup>. A 3.9-Å-resolution structure of IOMA–10-1074–BG505 was solved with four data sets collected from crystals prepared with early fractions from the BG505 SEC purification, and a 3.5-Å-resolution structure was solved with 14 data sets collected from crystals prepared with later BG505 fractions (Supplementary Fig. 4 and Table 1). The 3.9-Å-resolution structure was solved by molecular replacement with a monomeric gp120-gp41 from BG505 SOSIP (PDB 4TVF), one copy of 10-1074 Fab (PDB 4FQ2), and one copy of a chimeric Fab (PDB 4TNN for the LC and PDB 4XVS for the HC) identified as a suitable model for IOMA by the FFAS server<sup>55</sup>. Coordinates were refined with PHENIX v1.10.1-2155 (ref. 35) with group B factor and TLS restraints. The presence of ~10% twinning was corrected for by refining with the twin operator k,h,-l for space group R3:H. Manual rebuilding was performed with Coot<sup>56</sup>. The 3.5-Å-resolution structure was solved with a refined model from the 3.9-Å-resolution structure. Data refinement statistics are shown in Table 1. In both models, >90% of the residues were in the favored region of the Ramachandran plot, whereas <1% were in the disallowed regions.

Glycans were interpreted in both structures with 2F<sub>o</sub> – F<sub>c</sub> maps calculated with model phases and with composite annealed omit maps calculated with phases in which the model was omitted to reduce model bias<sup>35</sup>. Glycans were refined with PDB CArbohydrate REsidue check (pdbscare) (<http://www.glycosciences.de/tools/pdb-care/>), CArbohydrate Ramachandran Plot (carp) (<http://www.glycosciences.de/tools/carp/>), and Privateer<sup>37</sup>. In general, the glycans identified crystallographically at individual PNGSs on BG505 SOSIP.664 were found within the mixture of glycans assigned at these positions with mass spectroscopy<sup>11</sup> (Supplementary Figs. 5–8). However, the N-linked glycan attached to N156<sub>gp120</sub> was modeled as

complex type in our structures (**Supplementary Fig. 5**) but has been identified as high mannose (primarily  $\text{Man}_9\text{GlcNAc}_2$ ) in mass spectrometry analyses of BG505 SOSIP.664 (ref. 11). When we modeled a high-mannose glycan at this site and carried out refinement as described above, we found a slight increase (0.1%) in  $R_{\text{free}}$  and electron density in addition to that accounted for by the high-mannose residues, including density for a core fucose. We also modeled the N-linked glycan attached to N392<sub>gp120</sub> as containing a core fucose (**Supplementary Fig. 8**), but it has been identified as a  $\text{Man}_9\text{GlcNAc}_2/\text{Man}_8\text{GlcNAc}_2$  mixture in BG505 SOSIP.664 mass spectroscopy studies<sup>11</sup>. When the core fucose was removed from the N392<sub>gp120</sub> glycan coordinates, a subsequent refinement showed a minor increase in  $R_{\text{free}}$  and electron density for the fucose was preserved.

Buried surface areas were determined with PDBePISA<sup>57</sup> and a 1.4-Å probe. Superimposition calculations were performed, and molecular representations were generated with PyMOL (<http://www.pymol.org/>) or UCSF Chimera<sup>58</sup>. PDBeFold<sup>59</sup> was used to perform pairwise Cα alignments. The following distance and geometry criteria were used for assigning putative hydrogen bonds: a distance of <3.5 Å and an A-D-H angle of >90°. The maximum distance allowed for a van der Waals interaction of 4.0 Å. Hydrogen-bond and van der Waals interaction assignments should be considered tentative, owing to the relatively low resolutions of the structures.

**Antibody approach-angle comparisons.** The angles of approach for selected CD4bs bNAbs were compared as follows: the structure of CD4 in complex with HxBc2 gp120 (PDB **1GC1**) was used as a reference structure for comparisons of angles of approach of Fab recognition of gp120s. The center of mass of the CD4 D1 domain was placed at the origin, and its principal axes of inertia were aligned with the Cartesian axes with AMORE from the CCP4 suite<sup>54</sup>. The remainder of the PDB **1GC1** complex was then aligned with the centered CD4 D1 domain. To compare with other complexes, each Fab–gp120 complex was aligned with the **1GC1** gp120 chain with LSQMAN<sup>60</sup>. The transformation matrix between the aligned Fab–gp120 V<sub>H</sub> domain and the CD4 D1 domain was then calculated by LSQMAN.

**Antibody dendrogram calculation.** CD4-mimetic bNAb V<sub>H</sub>-gene-segment nucleotide sequences were aligned with Clustal Omega<sup>61</sup>, and the resulting

multiple sequence alignment was input into PhyML<sup>62</sup> to calculate a phylogenetic tree with tree topology and branch-length optimization (the tree improvement setting was ‘best of NNI and SPR’). Dendograms were drawn with FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

47. Murugan, R., Imkeller, K., Busse, C.E. & Wardemann, H. Direct high-throughput amplification and sequencing of immunoglobulin genes from single human B cells. *Eur. J. Immunol.* **45**, 2698–2700 (2015).
48. Tiller, T. *et al.* Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J. Immunol. Methods* **329**, 112–124 (2008).
49. von Boehmer, L. *et al.* Antigen-specific antibodies from mouse memory B cells. *Nat. Protoc.* (in the press) (2016).
50. Montefiori, D.C. Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. *Curr. Protoc. Immunol.* **64**, 12.11 (2005).
51. Sheskin, D. *Handbook of Parametric and Nonparametric Statistical Procedures* 3rd edn. (Chapman & Hall/CRC, 2004).
52. Galimidi, R.P. *et al.* Intra-spike crosslinking overcomes antibody evasion by HIV-1. *Cell* **160**, 433–446 (2015).
53. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 133–144 (2010).
54. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
55. Jaroszewski, L., Li, Z., Cai, X.H., Weber, C. & Godzik, A. FFAS server: novel features and applications. *Nucleic Acids Res.* **39**, W38–W44 (2011).
56. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
57. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
58. Pettersen, E.F. *et al.* UCSF Chimera: a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
59. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256–2268 (2004).
60. Kleywegt, G.J. Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr. D Biol. Crystallogr.* **52**, 842–857 (1996).
61. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
62. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).