

HIV-1 CD4-binding site germline antibody–Env structures inform vaccine design

Kim-Marie A. Dam¹, Christopher O. Barnes^{1‡}, Harry B. Gristick¹, Till Schoofs^{2,3,4§}, Michel C. Nussenzweig^{2,5}, Pamela J. Bjorkman¹

¹Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA

²Laboratory of Molecular Immunology, The Rockefeller University, New York, NY, USA

³Laboratory of Experimental Immunology, Institute of Virology, University of Cologne, Faculty of Medicine and University Hospital of Cologne, Cologne, Germany

⁴German Center for Infection Research, Partner Site Bonn–Cologne, Cologne, Germany

⁵Howard Hughes Medical Institute, Chevy Chase, MD, USA

[‡]Present address: Department of Biology, Stanford University, Stanford, CA, USA

[§]Present address: GlaxoSmithKline Vaccines, 1330 Rixensart, Belgium

Abstract

BG24, a VRC01-class broadly neutralizing antibody (bNAb) against HIV-1 Env with relatively few somatic hypermutations (SHMs), represents a promising target for vaccine strategies to elicit CD4-binding site (CD4bs) bNAbs. To understand how SHMs correlate with BG24 neutralization of HIV-1, we solved 4.1 Å and 3.4 Å single-particle cryo-EM structures of two inferred germline (iGL) BG24 precursors complexed with engineered Env-based immunogens lacking CD4bs N-glycans. Structures revealed critical Env contacts by BG24_{iGL} and identified antibody light chain structural features that impede Env recognition. In addition, biochemical data and cryo-EM structures of BG24_{iGL} variants bound to Envs with CD4bs glycans present provided insights into N-glycan accommodation, including structural modes of light chain adaptations in the presence of the N276_{gp120} glycan. Together, these findings revealed Env regions critical for germline antibody recognition and potential sites to alter in immunogen design.

Current strategies to engineer a vaccine towards preventing HIV-1 infection involve designing Env-mimetic immunogens that can elicit broadly neutralizing antibodies (bNAbs)¹⁻⁴. The CD4-binding site (CD4bs) epitope is a target of immunogen design as bNAbs in this class have been shown to be among the most potent and broad⁵⁻⁹. Several studies have shown passive immunization using CD4bs bNAbs can confer protection from HIV-1 infection in animal models and human clinical trials, suggesting that immunization strategies to elicit these antibodies at effective concentrations would also be protective^{6,10-17}. This includes the VRC01-class of bNAbs that are derived from the VH1-2*02 variable heavy chain gene segment and are characterized by a short 5 amino acid complementary determining region 3 (CDRL3) in the antibody (Ab) light chain and a shortened or flexible CDRL1^{5,18}. These characteristics are necessary for VRC01-class bNAbs to accommodate the heavily N-glycosylated landscape of the CD4bs of HIV-1 Envs. Thus, VRC01-class bNAbs generally require high levels of somatic hypermutation (SHM), which is challenging to elicit through vaccination.

Germline precursors of bNAbs do not generally show detectable binding to HIV-1 Envs^{19,20}, therefore, the germline targeting approach to HIV-1 vaccine design involves efforts to engineer immunogens that can engage germline B cell receptors (BCRs) and initiate bNAb development²¹. Inferred germline (iGL) versions of mature bNAbs derived from predicted germline gene segment sequences represented in the human B cell repertoire^{22,23} are used for the germline targeting approach. Analysis of VRC01-class iGLs has shown that the human VH1-2*02 heavy chain gene segment encodes signature residues that are required for breadth and potency¹⁸. Furthermore, germline VRC01-class precursors have been isolated from naïve individuals and mature bNAbs have been identified from multiple HIV-1-infected human donors, suggesting that raising this class of bNAbs is not uncommon in natural infection^{24,25}. Taken together, VRC01-class bNAbs are attractive targets for immunogen design.

The VRC01-class of bNAbs target the CD4bs, a particularly challenging epitope to elicit bNAbs against due to the presence of the CD4bs N-glycans that sterically obstruct interactions between Env and Ab CDRs²⁶. The glycan at position N276_{gp120} is highly conserved and poses the greatest steric barrier to binding VRC01-class bNAb iGLs, as Ab residues in the iGL CDRL1 that interact with this region are typically 11-12 residues and cannot accommodate the N276_{gp120} glycan. Mature CD4bs Abs develop shortened or flexible CDRL1s to accommodate this glycan^{24,27,28}. Thus, understanding the structural basis for how CD4bs iGL Abs mature to effectively accommodate the N276_{gp120} glycan is essential in efforts to develop effective immunogens to prime VRC01-class iGL precursors and shepherd antibody responses towards bNAb development. Furthermore, an overall structural understanding of VRC01-class iGL recognition of HIV-1 Envs and immunogens is limited as the only existing Fab-Env structures involving germline CD4bs Abs are complexed with gp120 or Env trimer immunogens lacking the N276_{gp120} glycan^{3,23,29}. In addition, in the case of an iGL Fab complexed with an Env trimer, obtaining a structure required chemical cross-linking between the Env and Ab to form a stable complex²².

A new VRC01-class bNAb isolated from an elite neutralizer, BG24³⁰, is an attractive target for germline-targeting immunogen design. BG24 shows similar neutralization and breadth to other CD4bs bNAbs, but includes only 22.6% and 19.5% amino acid substitution by SHM in variable heavy and light chain genes, respectively³⁰, as compared with higher levels of amino acid substitution in VRC01-class bNAbs^{7,9,28,31}, with the exception of the PCIN63 lineage that has similar levels of SHM to BG24³². Structural characterization of BG24 bound to the clade A BG505 Env revealed a similar binding orientation to more mutated VRC01-class bNAbs, and signature contacts common to VRC01-class bNAbs³⁰. Furthermore, neutralization studies using variants of BG24 that reverted variable heavy (V_H) and variable light (V_L) domain residues to germline counterparts showed that even fewer SHMs were necessary to maintain neutralization

breadth³⁰. Collectively, this suggests broad and potent neutralization targeting the CD4bs could be achieved through immunization without stimulating high levels of SHM.

To better understand how the BG24 bNAbs was elicited and inform VRC01-class immunogen design, we structurally characterized the binding of two versions of the BG24 iGL to the CD4bs germline-targeting immunogen BG505-SOSIPv4.1-GT1³ (hereafter referred to as GT1). We solved two single-particle cryo-electron microscopy (cryo-EM) structures of GT1 in complex with BG24_{iGL}s containing either mature or iGL CDR3s at 4.1 Å and 3.4 Å resolution, respectively, in both cases in the absence of chemical cross-linking. Furthermore, to understand how N-glycans impact germline Ab recognition of Env, we conducted biochemical assays and solved cryo-EM structures of BG24_{iGL} derivatives bound to Envs that included the N276_{gp120} glycan. The structures demonstrated that the CDRL1s of BG24_{iGL}s can adopt conformations that accommodate the N276_{gp120} glycan, an important capability for a germline-targeting CD4bs immunogen. Collectively, these structures provide information regarding the physical characteristics of iGLs that recognize HIV-1 Env and provide a structural basis for the design of immunogens engineered to engage and mature germline Abs.

Results

Cryo-EM structures of GT1-BG24_{iGL}-101074 complexes

To gain insight into how BG24 precursors interact with an HIV-1 Env-based immunogen, we created iGL versions of BG24 and used single-particle cryo-EM to structurally characterize them in complex with GT1, a CD4bs germline-targeting immunogen³. GT1 was modified from a soluble clade A BG505 SOSIP.664 native-like Env trimer³³ to permit binding of VRC01-class germline precursors by including T278R_{gp120} and G271S_{gp120} substitutions and mutations to remove potential N-linked glycosylation sites (PNGSs) at positions N276_{gp120}, N462_{gp120}, N386_{gp120}, and N197_{gp120} in the CD4bs³. Two iGL versions of BG24 Fab constructs were made

starting with the VH1-2*02 and VL2-11*01 heavy and light chain germline gene segment sequences: one containing the CDR3s from mature BG24 (BG24_{iGL-CDR3mat}) and the other containing the iGL CDR3s (BG24_{iGL-CDR3iGL}) (Fig. 1a). Each BG24_{iGL} was structurally characterized in complex with GT1 and the V3 bNAbs 10-1074³⁴.

Cryo-EM structures of BG24_{iGL-CDR3iGL} and BG24_{iGL-CDR3mat} Fabs bound to GT1 were solved at 3.4 Å and 4.1 Å, respectively (Fig. 1b,c, Extended Data Fig. 1a-j, Supplementary Table 1). Both 3D cryo-EM reconstructions showed three BG24_{iGL} and three 10-1074 Fabs bound per Env trimer. However, for the BG24_{iGL-CDR3iGL}-GT1-101074 complex, a distinct 3D class contained two BG24_{iGL-CDR3iGL} Fabs bound to the GT1 Env (Extended Data Fig. 1e-f, Supplementary Table 1). We also solved a 1.4 Å crystal structure of unbound BG24_{iGL-CDR3mat} Fab (Extended Data Fig. 1k, Supplementary Table 2), which exhibited six disordered residues within CDRL1, but otherwise superimposed with a 1.3 Å root mean square deviation (rmsd; calculated for 225 V_H-V_L C α atoms) with the Env-bound BG24_{iGL-CDR3mat} Fab structure, suggesting no major structural differences upon Env binding.

BG24_{iGL} Fabs recognize the modified CD4bs in GT1 Env

The GT1 complexes with BG24_{iGL}s included density for CD4bs N-glycans attached to residues N234_{gp120}, N363_{gp120}, and N392_{gp120} (Fig. 2a-b). These N-glycans were also observed in the crystal structure of BG505 Env complexed with a mature BG24 Fab³⁰ (BG24_{mat}) (PDB 7UCE), which also included densities for N-glycans at N197_{gp120}, N276_{gp120}, and N386_{gp120} that are not present in GT1 (Fig. 2c). Despite additional glycans in BG505 compared with GT1, the CDR loops in the GT1-bound iGL Fabs showed similar orientations and positions as in the BG505-bound BG24_{mat} Fab, except for CDRL1, which is six residues longer in BG24_{iGL} than in BG24_{mat} (Fig. 1a, 2d-f).

BG24_{iGL-CDR3mat} and BG24_{iGL-CDR3iGL} buried comparable surface areas on GT1 gp120 (960 Å² and 951 Å², respectively) as compared with an only slightly larger surface area (1099 Å²) buried on BG505 gp120 in the BG24_{mat}-BG505 structure (PDB 7UCF) (Fig. 2a-b, g). However, differences in interactions between the BG24_{iGL}-GT1 and BG24_{mat}-BG505 structures suggest that SHM substitutions enrich interactions in particular regions within the CD4bs. For example, in the BG24_{mat}-BG505 complex, BG24_{mat} residue S99_{HC} hydrogen bonds with the gp120 inner domain residue K97_{gp120} (Fig. 2c, 2g). K97_{gp120} is ~90% conserved among HIV-1 Envs, making this a crucial interaction of broad and potent CD4bs bNAb¹⁸. Residue S99_{HC} is a germline encoded residue, however, in both BG24_{iGL}-GT1 structures, it did not hydrogen bond with K97_{gp120}. Compared to BG24_{iGL}-GT1, BG24_{mat}-BG505 also showed increased V_H buried surface area (BSA) in the gp120 exit loop (gp120 residues 470-474), and gained BSA in the gp120 alpha 5 region, suggesting BG24_{mat} has a larger BSA footprint than the BG24_{iGL}s in these CD4bs regions (Fig. 2c, 2g). Interestingly, the overall interface BSA values for the gp120 peptide components of the BG24_{iGL}-GT1 and BG24_{mat}-BG505 structures were similar (960 and 951 Å² vs 1099 Å², respectively). Although a germline precursor antibody presumably exhibits fewer contacts to an antigen than its counterpart somatically-mutated bNAb, the modifications in GT1 (both amino acid substitutions and removals of N-glycans) allowed increased contacts between BG24_{iGL}s and the GT1 gp120.

BG24 somatic hypermutation plays a role in CD4bs recognition

We next compared how differences in BG24_{iGL} and BG24_{mat} contribute to their recognition of GT1 and BG505, respectively. BG24_{iGL} contains a germline 11-residue CDRL1 that can recognize the mostly aglycosylated CD4bs in GT1, whereas the BG24_{mat} CDRL1 is six residues shorter and includes a glycine to create a more flexible loop that can accommodate the N276_{gp120} glycan³⁰. In the BG24_{mat}-BG505 structure, the five-residue BG24 CDRL1 is oriented adjacent to the N276_{gp120} glycan (Fig. 3a). The CDRL1 interface with GT1 in the BG24_{iGL-CDR3iGL}

and BG24_{iGL-CDR3mat} structures showed the longer CDRL1s in the germline precursor V_L domains in different conformations, demonstrating CDRL1 flexibility (Fig. 3b-c) consistent with cryo-EM data processing. Indeed, the local resolutions for the CDRL1 in these structures were poor and resolved only after iterative rounds of focused classification and local refinements (Extended Data Fig. 2a-b). Overlaying the BG24_{iGL} CDRL1s with the gp120 region surrounding the N276_{gp120} glycan from the BG24_{mat}-BG505 structure showed steric clashes, consistent with SHM being necessary for N276_{gp120} glycan accommodation by BG24 (Fig. 3b-c).

The role of SHMs in Env recognition is summarized in Fig. 3d, where BG24_{iGL-CDR3iGL}, BG24_{iGL-CDR3mat}, and BG24_{mat} HC paratope interactions are mapped to individual Ab residues within 4 Å of gp120. Paratope contacts were limited to CDRs H1, H2, and H3, as well as framework region 3 in the heavy chain (FWRH3), with most contacts in CDRH2. Previous studies showed neutralization by an engineered BG24 “minimal” construct that contained germline-reverted SHMs in FWRs, CDRH1, and CDRL2, but maintained most SHMs in CDRH2, suggesting the importance of SHMs in this region³⁰. The structure of BG24_{mat}-BG505 showed a CDRH2 SHM (N53_{R_{HC}}) interacted with Q428_{gp120} in gp120 β20/21 (Fig. 1a, 3e). β20/21 interactions with germline-encoded N53_{HC} were absent in BG24_{iGL-CDR3iGL}-GT1 and BG24_{iGL-CDR3mat}-GT1 (Fig. 3f-g). This demonstrates the direct impact of SHM in creating favorable interactions with Env. Other BG24_{mat} somatically hypermutated residues in CDRH2 also interacted with the CD4bs loop (gp120 residues 356-371); e.g., residue T57_{V_{HC}} makes a backbone interaction with S365_{gp120}, and S54_{G_{HC}} interacts with D368_{gp120}, a highly conserved Env residue (Fig. 3d)¹⁸. For BG24_{iGL}, germline-encoded residues at positions T57_{HC} and S54_{HC}, maintain similar interactions with GT1 residues S365_{gp120} and D368_{gp120}, respectively, suggesting that these germline encoded residues may be sufficient to engage the CD4bs loop (Fig. 3f-g).

Signature residues encoded by the VH1-2*02 germline gene in VRC01-class bNAbs interact with conserved gp120 residues and are correlated with neutralization potency¹⁸. These interactions have been structurally characterized in the context of VRC01-class iGLs bound to monomeric gp120s^{22,23,29}, but there are no known structures of VRC01-class iGLs bound to a trimeric Env, except when the iGL was chemically cross-linked to Env²². To evaluate and verify VRC01-class VH1-2*02 germline-encoded interactions with an Env trimer, we compared these interactions in the BG24_{iGL}-GT1 and BG24_{mat}-BG505 structures (Extended Data Fig. 2c-f). Specifically, as previously described in structures involving gp120 monomers^{22,23,29}, germline-encoded R71_{HC} in the BG24_{iGL}-GT1 and BG24_{mat}-BG505 structures formed a salt bridge with the conserved D368_{gp120} side chain, an Ab interaction that mimics the interaction of host receptor residue R59_{CD4} with D368_{gp120} (Extended Data Fig. 2c). In the gp120 D loop, there were interactions between the backbone and side chain of N280_{gp120} with Y100_{B_{HC}} and germline-encoded W50_{HC} side chains (Extended Data Fig. 2d). In the V5 loop, interactions between the conserved R456_{gp120} residue and germline-encoded N58_{HC} are conserved in both structures (Extended Data Fig. 2e). In BG24_{iGL} CDR3_{mat}-GT1, atoms within these Fab residues were separated by more than 5 Å from atoms within gp120 residues; thus, this is not defined as an interaction. In the light chain, E96_{LC} interacted with the backbone of G456_{gp120} and the sidechain of N280_{gp120} (Extended Data Fig. 2f).

GT1 CD4bs glycan modifications affect BG24 binding

To evaluate how glycan modifications in the GT1 immunogen contributed to BG24_{iGL} binding, we evaluated the binding of BG24 constructs to GT1 with Env PNGSs either restored to or removed from the CD4bs (Fig. 4a). The BG24 constructs included BG24_{mat}, BG24 with germline CDRL1 (BG24_{CDRL1-iGL}) (Fig. 1a), BG24_{iGL-CDR3mat}, BG24_{iGL-CDR3iGL}, and BG24 with an iGL light chain (BG24_{LC-iGL}). PNGSs were individually restored at positions N197_{gp120}, N276_{gp120}, N386_{gp120}, and N462_{gp120} and removed at N234_{gp120} to create five GT1 constructs with altered glycan

landscapes (GT1_{N197gp120}, GT1_{N276gp120}, GT1_{N386gp120}, GT1_{N462gp120}, and GT1_{del N234gp120}, respectively). BG505 and GT1 binding was evaluated by enzyme-linked immunosorbent assays (ELISAs). Restoring Env PNGSs at positions N197_{gp120}, N386_{gp120}, and N462_{gp120} and removing the PNGS at N234_{gp120} did not greatly affect binding of BG24 IgG constructs (Fig. 4a). BG24_{iGL} constructs did not bind detectably to GT1 with a PNGS at N276_{gp120}; however, BG24_{mat} and mature BG24 constructs with iGL LC features (BG24_{CDRL1-iGL}, BG24_{LC-iGL}) showed comparable binding to each other on GT1_{N276gp120} (Fig 4a).” BG24_{mat} was the only Ab that showed substantial binding to BG505, which unlike the GT1 Env, included all PNGSs. We conclude that BG24 constructs with a long, germline CDRL1 can accommodate the N276_{gp120} glycan on Envs that have been engineered to have a limited glycan landscape in the CD4bs.

To gain insight into BG24 CDRL1-iGL interactions in GT1 with an N276_{gp120} glycan, we solved a single-particle cryo-EM structure of BG24_{LC-iGL} bound to GT1 containing the restored N276_{gp120} PNGS (GT1_{N276gp120}) (Fig. 4b, Extended Data Fig. 3a-h, Supplementary Table 1). We identified three unique 3D volumes containing either one, two, or three bound BG24_{LC-iGL} Fabs, with the highest resolution complex (4.2 Å) being C3 symmetric with three bound BG24_{LC-iGL} Fabs (Extended Data Fig. 3a-h, Supplementary Table 1). Electron density in the Fab CDRL1 was not optimal after cryo-EM processing; therefore, side chains were not modeled (Extended Data Fig. 3i).

The BG24_{LC-iGL}-GT1_{N276gp120} complex structure showed that the Fab CDRL1 main chain residues adopted a helix-like conformation to accommodate the N276_{gp120} glycan (Fig. 4c). Available crystallographic and cryo-EM Env structures demonstrate that the N276_{gp120} glycan is conformationally heterogeneous (Extended Data Fig. 3h). Indeed, the N276_{gp120} glycans in the GT1_{N276gp120} and BG505 Envs exhibited different conformations (Fig. 4c). Thus, after superimposing the gp120 residues in the BG24_{LC-iGL}-GT1_{N276gp120} and BG24_{mat}-BG505

structures, it was evident that the N276_{gp120} glycan in BG505 showed steric clashes with the CDRL1 iGL in BG24_{LC-iGL} (Fig. 4c). Flexibility of the N276_{gp120} glycan on BG505 may be more constrained than the counterpart glycan on GT1, as GT1 contains fewer N-glycans in the CD4bs, allowing for increased N276_{gp120} glycan flexibility. This assumption is consistent with the ELISA results showing that BG24_{LC-iGL} bound to GT1, but not to BG505 Env trimers with an N276_{gp120} glycan (Fig. 4a).

The only other known CD4bs-targeting bNAb with a helical CDRL1 is IOMA, another VH1-2*02 derived bNAb²⁷. IOMA contains features that distinguish it from VRC01-class bNAbs, including a normal length (8 residue) CDRL3 and a 13-residue CDRL1, which adopts a short α -helix to accommodate the glycan at N276_{gp120}. However, CLK31, an IOMA-like Ab isolated from naïve human B cells using a VRC01 germline-targeting immunogen, did not include a helical CDRL1³⁵. Alignment of the LCs of BG24_{LC-iGL}, IOMA, and CLK31 showed that each CDRL1 adopts a different configuration (Fig. 4d). These observations suggest CDRL1 helical conformations are diverse and have only been observed when bound to gp120s that contain the glycan at N276_{gp120}.

BG24_{CDRL1-iGL} accommodates the N276_{gp120} glycan in a non-engineered Env trimer

A longitudinal study that tracked the development of a VRC01-class lineage (PCIN63) found that bNAb development branched into two types of N276_{gp120} glycan engagement: one that interacted with and depended on the presence of the N276_{gp120} glycan, and one in which CD4bs binding was diminished by the presence of the N276_{gp120} glycan³². In the absence of longitudinal data for BG24_{mat} development, a BG24 intermediate, BG24_{CDRL1-iGL}, was tested for neutralization against a 119-virus cross clade panel to better understand how the germline BG24 CDRL1 interacted with HIV-1 Envs bearing the N276_{gp120} glycan. BG24_{CDRL1-iGL} exhibited neutralization activity against two viruses that contained PNGSs at N276_{gp120}: clade D 6405.v4.c34 (6405) and

clade CD 6952.v1.c20 (6952) (Fig. 5a). The 6405 Env was selected for further investigation by creating a soluble 6405 SOSIP.664 trimer. Sequence alignment of 6405 and BG505 gp120s showed the amino acid identity in the CD4bs and V4 loops differed by more than 50% between BG505 and 6405 Envs (Extended Data Fig. 4). The 6405 gp120 sequence included similar CD4bs PNGSs as BG505, except for the absence of a PNGS at position 363_{gp120} and an added PNGS at position 465_{gp120} (Extended Data Fig. 4). Binding of BG24 Fab constructs to 6405 was assessed by ELISA. Consistent with neutralization results (Fig. 5a), ELISAs showed that BG24_{CDRL1-iGL} and BG24_{mat} each bound the 6405 SOSIP, whereas BG24_{LC-iGL} bound 6405 to a lesser extent (Fig. 5b). BG24_{iGLS} did not bind detectably to 6405 (Fig. 5b).

To understand how the germline CDRL1 of BG24 could be accommodated by a non-engineered Env trimer, we characterized interactions between BG24_{CDRL1-iGL} and 6405 SOSIP by solving a 3.4 Å cryo-EM structure of a BG24_{CDRL1-iGL}-6405 complex (Fig. 5c, Extended Data Fig. 5a-d, Supplementary Table 1). As expected, BG24_{CDRL1-iGL} recognized the CD4bs of 6405, which contained N-glycans at positions N197_{gp120}, N234_{gp120}, N276_{gp120}, N354_{gp120}, N386_{gp120}, N392_{gp120}, N461_{gp120}, and N465_{gp120} (Fig. 5d). Again, side chains were not modeled for CDRL1 residues (Extended Data Fig. 5e). As also observed in the BG24_{LC-iGL}-GT1_{N276} structure, the CDRL1-iGL in the BG24_{CDRL1-iGL}-6405 complex formed a helical conformation, although the CDRL1 conformations in the two Fabs in these complexes were not identical (Extended Data Fig. 5f). By aligning gp120s from GT1_{N276} and 6405, we found that the well-ordered portions of the N276_{gp120} glycan occupied similar positions (Fig. 5f), suggesting that 6405 Env contains a native glycan landscape conducive to glycan flexibility at position N276_{gp120}, similar to what was observed in this region of the GT1 Env structure.

We also evaluated binding of other CD4bs bNAbs to 6405 and a 6405_{delN276gp120} Env to determine if the CD4bs glycan landscape in 6405 was conducive to interactions with germline

CDRL1s in other VRC01-class bNAbs (Fig. 5f-g). In this experiment, we included mature versions of the BG24, VRC01, N6 and IOMA bNAb Fabs, chimeric bNAb Fabs including an iGL LC, and complete iGL Fabs. PCIN63 and CH103 and intermediates were compared with their unmutated common ancestors (UCAs), as identified from longitudinal studies^{32,36}, instead of iGLs. The ELISA revealed that the 6405 Env interacted with VRC01_{LC-iGL} and N6_{LC-iGL} in addition to BG24_{LC-iGL}. Binding for all species increased for 6405_{del N276gp120}, indicating the N276_{gp120} glycan can sterically impede the CDRL1 iGL despite its flexibility. We conclude that the 6405 Env contains a CD4bs glycan landscape that tolerates the binding of germline VRC01-class CDRL1s.

Discussion

VRC01-class bNAbs are promising targets for germline-targeting immunogen design as germline-encoded residues make signature contacts with gp120 that contribute to impressive breadth and potency^{5,18,22,23}. However, challenges in eliciting VRC01-class bNAbs through a germline-targeting vaccine regimen include explicitly selecting for the VH1-2*02 germline gene, overcoming CD4bs glycan barriers, and stimulating high levels of SHM^{37,38}. Overcoming these challenges will be facilitated by increased understanding of how germline VRC01-class Abs interact with germline-targeting immunogens.

Despite these challenges, progress has been made in developing a VRC01-class bNAb germline-targeting approach^{21,38}, which is initiated by engineering an immunogen that binds to the germline precursors of VRC01 bNAbs. Priming immunogens are engineered to interact with specific germline-encoded residues and lack CD4bs glycans that obstruct germline recognition^{3,29,37,39}. VRC01-class priming immunogens include monomeric gp120 cores^{25,37,40}, SOSIP-based trimers³, and anti-idiotypic antibodies that recognize target BCRs with VH1-2*02 gene segments^{41,42}. Selecting a particular strain of Env and a gp120- versus trimeric Env-based

platform to engineer priming characteristics is important. For example, a recent study comparing three different gp120-based immunogens found that the epitope and surrounding surface of the immunogen affects germline BCR selection in vivo⁴³. Thus, identifying and developing the optimal priming immunogen for VRC01-like bNAb elicitation will require a robust understanding of the structural and biophysical nature of Env recognition by germline precursors.

In a sequential immunization approach, boosting immunogens are introduced to shape the development of a germline precursor into a bNAb by stimulating favorable SHMs³⁸. Example boosting immunogens re-introduce native Env glycans and heterogenous Env strains to develop bNAbs capable of overcoming steric glycan barriers and have heterologous-neutralizing activity³⁸. The N276_{gp120} glycan on HIV-1 Env provides a particularly difficult roadblock, as VRC01-like germline CDRL1s must become shorter or more flexible through SHM to avoid steric clashes^{23,27,28,32,44}. Several iterations of this approach have been tested in animal models; however, the elicitation of heterologous neutralizing activity has not yet been accomplished^{37,45}.

BG24_{mat} represents a new VRC01-class bNAb that can be targeted for germline-targeting approaches³⁰. BG24_{mat} has a fraction of the SHMs found in VRC01 and other VRC01-class bNAbs and maintains notable breadth and potency. Together with previous studies of the VRC01-class PCIN63 lineage and construction of a minimally mutated VRC01, our studies of BG24 suggest that high levels of SHM are not absolutely required for the development of VRC01-class Abs^{32,44}. Our cryo-EM structures of the iGL precursors of BG24 bound to the priming immunogen GT1 contribute to understanding how VRC01-class bNAb precursors interact with immunogens. We found that VH1-2*02 germline-encoded residues make the predicted signature contacts with gp120 and the long germline CDRL1 is accommodated in the absence of the N276_{gp120} glycan in GT1, rationalizing removal of this glycan in a priming immunogen since modeling suggested the germline CDRL1 conformation would clash with the

N276_{gp120} glycan. These observations validate the design of priming immunogens that nurture interactions with germline residues and remove the N276_{gp120} glycan from the CD4bs epitope. We further investigated how the glycan landscape of an immunogen affects germline binding, finding that BG24_{iGL-LC} can evade clashes with the N276_{gp120} glycan when the BG24 HC includes bNAbs features and the CD4bs epitope is only minimally glycosylated. Based on these observations, we propose that boosting immunogens might first aim to target mature HC features, and then introduce the N276_{gp120} glycan in a limited CD4bs glycan landscape before moving to fully glycosylated Env landscape.

We also characterized binding of BG24_{CDRL1-iGL} to the clade D 6405 Env, which suggested that some non-engineered HIV-1 Envs contain glycan landscapes that can accommodate germline VRC01-class CDRL1s. In the case of BG24_{CDRL1-iGL}, accommodation of the N276_{gp120} glycan occurred through a helix-like conformation in the iGL CDRL1. Furthermore, ELISA data suggested that other VH1-2*02-derived bNAbs with iGL LCs can also bind to the 6405 Env. Taken together, we believe the 6405 Env contains properties that make it a desirable candidate for a boosting immunogen to introduce the N276_{gp120} glycan after priming: in particular, unlike most CD4bs-targeting immunogens used as early boosting immunogens, 6405 is a non-engineered Env that can interact productively with germline LCs. Therefore, we propose further investigation of 6405 Env in boosting regimens to determine if these structural and biochemical observations translate in vivo.

Methods

BG24_{iGL} constructs design

Genes encoding the IGHV1-2*02 and IGLV2-11*02 germline sequences with mature CDR3 loops were used to generate the BG24_{iGL-CDR3mat} Fab construct. For the BG24_{iGL-CDR3iGL} construct, amino acids in D- and J- gene regions were reverted based on inferred sequences using IMGT.

Protein expression and purification

Fabs and IgGs were expressed and purified as previously described⁴⁶. Briefly, Fabs were expressed by transient transfection using the Expi293 expression system (ThermoFisher). Fab expression vectors contained genes of LC and the C-terminally 6x-His tagged HC. The Fab and IgG proteins were purified from cell supernatants by Ni²⁺-NTA (GE Healthcare) and protein A affinity chromatography (GE Healthcare), respectively, followed by size exclusion chromatography (SEC) using a Superdex 200 10/300 column (GE Healthcare).

SOSIP.664 Env constructs contained the disulfide mutations 501C and 650C (SOS), I55P (IP), and the furin cleavage site mutated to six arginine residues (6R)³³. Genes encoding BG505 SOSIP.664v4.1-GT1 and 6405 SOSIPs were expressed using by transient transfection of Expi293 cells (ThermoFisher) and purified as described previously⁴⁷. The 6405 SOSIP construct contained gp120 residues 46-477 from the 6405 sequence, with remaining gp120 residues derived from BG505 and the extracellular portion of the BG505 gp41⁴⁸. Trimeric Env was separated from cell supernatants by PGT145 immunoaffinity chromatography, and SEC using a Superose 6 10/300 column (GE Healthcare) as described⁴⁹.

X-ray crystallography

Purified BG24_{iGL-CDR3mat} Fab was concentrated to 8-15 mg/mL. Matrix crystallization screens were performed at room temperature using the sitting drop vapor diffusion method by mixing equal volumes of protein sample and reservoir using a TTP LabTech Mosquito robot and commercially-available screens (Hampton Research and Qiagen). Initial hits were optimized and crystals were obtained in 20% PEG 3350 at 20 °C. Crystals were cryo-protected in glycerol stepwise until 20% before being cryopreserved in liquid nitrogen.

X-ray diffraction data were collected to 1.4 Å for BG24_{iGL-CDR3mat} Fab at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 12-2 on a Pilatus 6M pixel detector (Dectris). Data from a single crystal were indexed and integrated in XDS⁵⁰ and merged with AIMLESS in the CCP4 software suite⁵¹. Structures were determined by molecular replacement in PHASER⁵² using coordinates of the BG24_{mat} Fab (PDB 7UCE), after removal of CDR loops and independent searches of the V_HV_L and C_HC_L domains. Models were refined using rigid body and B-factor refinement in Phenix⁵³, followed by several cycles of iterative manual building in Coot⁵⁴ and real-space refinement with TLS groups in Phenix^{53,55} (Supplementary Table 2).

Assembly of protein complexes and cryo-EM sample preparation

Protein complexes for cryo-EM were generated by incubating a purified BG24_{iGL} Fab and the 10-1074 Fab with an Env trimer in a 3:1 Fab:trimer molar ratio and incubating at 4°C overnight. The complex was then SEC purified over a Superdex 200 1/150 column (GE Healthcare). The peak corresponding to complex was pooled and concentrated to 1.0 mg/ml. Quantifoil R2/2 400 mesh cryo-EM grids (Ted Pella) were prepared by glow-discharging for 1 min at 20 mA using a PELCO easiGLOW (Ted Pella). Fab-Env complexes (3 µL) were then applied to grids and blotted with Whatman No. 1 filter paper for 3-4 s at 100% humidity at room temperature. The

grids were vitrified by and plunge-freezing in liquid ethane using a Mark IV Vitrobot (ThermoFisher).

Cryo-EM data collection and processing

Data for single-particle cryo-EM were collected on either a Talos Artica (BG24_{iGL-CDR3mat}-GT1-10-1074, BG24_{iGL-CDR3iGL}-GT1_{N276gp120}-10-1074, BG24_{CDRL1-iGL}-6405-10-1074) or a Titan Krios (BG24_{iGL-CDR3mat}-GT1-10-1074) transmission electron microscope, operating at 200 kV and 300 kV, respectively. Movies were collected with beam-image shift over a single exposure per hole in a 3-by-3 pattern of 2 μm holes. For datasets collect on the Talos Artica, movies were recorded in super-resolution mode on a Falcon III camera (Thermo Fisher) at $1.436 \text{ \AA}\cdot\text{pixel}^{-1}$ or a K3 camera (Gatan) at $0.4345 \text{ \AA}\cdot\text{pixel}^{-1}$. Movies obtained from samples on the Titan Krios were collected in super-resolution mode on a K3 camera (Gatan) equipped with an BioQuantum energy filter (Gatan) with a 20 eV slit width at $0.4327 \text{ \AA}\cdot\text{pixel}^{-1}$. The defocus range was set from 1.0-3.0 μm for each dataset.

The data processing workflow described below was preformed similarly for all datasets using RELION^{56,57}. Movies were motion-corrected using MotionCor2⁵⁸ after binning. GCTF⁵⁹ was used to estimate CTF, and micrograph power spectra that showed poor CTF fits or bad ice were removed. A subset of particles was manually picked and used for reference-free 2D classification. Classes representing the defined complex were used as references for RELION AutoPicking^{56,57} to generate 2D classes. Subsequent 2D classes were inspected, and 2D classes representing a defined complex were selected for 3D classification. An ab initio model was generated using cryoSPARC⁶⁰ using a subset of particles for each dataset and used as a reference in 3D classification which assumed C1 symmetry. 3D classes representing a defined complex were selected for 3D auto-refinement and post processing in Relion. Particles used in 3D refinement were then re-extracted and un-binned. Particles were then subjected to 3D

classification with the map generated with un-binned particles used as a reference. Distinct classes representing a particular defined complex (C1 or C3 symmetric) were selected for 3D auto-refinement after masking out Fab C_HC_L domains. Iterative rounds of particle CTF refinement, particle polishing, 3D auto-refinement, and post processing were used for each class to generate final maps. To improve resolution of Fab_{LC} CDRL1s, a soft mask surrounding the Fab VH-VL-gp120 interface was created in chimera and used for local refinements in cryoSPARC to improve density in this region and allow for CDRL1 fitting and refinement. Resolutions were calculated in RELION using the gold-standard FSC 0.143 criterion⁶¹. FSCs were generated by the 3DFSC program⁶².

Cryo-EM model building and refinement

Model coordinates were generated by fitting reference gp120 (PDB 5T3Z), gp41 (PDB 5T3Z), 10-1074 (PDB 5T3Z), and BG24-derivative Fabs (this study) chains into cryo-EM density with UCSF Chimera⁶³. Initial models were refined using the Phenix command *phenix.real_space_refine*^{53,55}. Sequence updates to the model and further manual refinement was conducted with Coot⁵⁴. Iterative rounds of Phenix auto-refinement and manual refinements were done to generate the final models (Supplementary Table 1).

Structural analyses

Structure figures were made using PyMol (Schrödinger LLC), UCSF Chimera⁶³, and UCSF ChimeraX^{64,65}. PyMol was used to calculate r.m.s.d. values after pairwise alignment of Ca atoms. PDBePISA⁶⁶ was used to calculate buried surface areas using a 1.4 Å probe. Calculations for gp120 BSA were for peptide components of gp120 and did not include glycan interactions. Defined interactions were assigned tentatively due to the low resolution of complexes using the following criteria: hydrogen bonds were assigned pairwise interactions that

were less than 4.0 Å and with an A-D-H angle >90°, and van der Waals interactions were assigned as distances between atoms that were less than 4.0 Å.

Enzyme-linked immunosorbent assay

SOSIP trimers were randomly biotinylated following manufacturer's guidelines using the EZ-Link NHS-PEG4-Biotin kit (Thermo Fisher Scientific). The Pierce Biotin kit (Thermo Fisher Scientific) was used to quantify biotin molecules per SOSIP protomer: biotin estimations ranged from 1-10 biotin molecules per SOSIP protomers. Streptavidin-coated 96-well plates (Thermo Fisher Scientific) were coated with 5 µg/mL of randomly biotinylated SOSIPs diluted in 3% BSA in TBS-T (20mM Tris, 150 mM NaCl, 0.1% Tween20) and incubated at room temperature (RT) for 2 hours. Plates were washed to remove unbound SOSIPs. Serial dilutions of IgGs were made in 3% BSA in TBS-T and applied to the plates. After a 2-hour incubation at RT, plates were washed twice in TBS-T. Goat anti-human IgG Fc conjugated to horse-radish peroxidase (Southern BioTech,) was added at 1:8000 dilution for 30 minutes, followed by 3 washes with TBS-T. 1-Step™ Ultra TMB-ELISA Substrate Solution (ThermoFisher Scientific) was added for colorimetric detection and color development was quenched with 1N HCl. Absorbance was measured at 450nm. Two independent, biological replicates were performed.

Acknowledgements

We thank J. Vielmetter, P. Hoffman, and the Protein Expression Center in the Beckman Institute at Caltech for expression assistance. Electron microscopy was performed in the Caltech Cryo-EM Center with assistance from S. Chen and A. Malyutin. We thank the Gordon and Betty Moore and Beckman Foundations for gifts to Caltech to support the Molecular Observatory (Dr. Jens Kaiser, Director) and the beamline staff at SSRL for data collection assistance. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is

supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-c76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (P41GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH. C.O.B. was supported by a Burroughs Wellcome PDEP fellowship and HHMI Hanna Gray Fellowship. This work was supported by the National Institute of Allergy and Infectious Diseases (NIAID) Grant HIVRAD P01 AI100148 (to P.J.B. and M.C.N.), the Bill and Melinda Gates Foundation Collaboration for AIDS Vaccine Discovery (CAVD) grant INV-002143 (P.J.B., M.C.N.), and NIH P50 AI150464 (P.J.B.).

Author contributions

K.A.D., C.O.B., H.B.G., T.S., M.C.N., and P.J.B. designed the research. K.A.D. performed experiments and K.A.D., C.O.B., and H.B.G. analyzed results. K.A.D. and P.J.B. wrote the manuscript with input from co-authors.

Competing interests

The authors declare that there are no competing interests.

Data availability

The atomic model generated for the X-ray crystallography structure of the BG24_{iGL-CDR3mat} Fab in this study has been deposited in the Protein Data Bank (PDB) under accession code 7UGM. The cryo-EM maps and atomic structures have been deposited in the PDB and/or Electron Microscopy Data Bank (EMDB) under accession codes 7UGN and EMD-26490 for BG24_{iGL-CDR3iGL-GT1-10-1074 Class 1}, EMD-26491 for BG24_{iGL-CDR3iGL-GT1-10-1074 Class 2},

7UGO and EMD-26492 for BG24_{iGL-CDR3mat}-GT1-10-1074, 7UGP and EMD-26493 for BG24_{iGL-LC}-GT1_{N276gp120}-10-1074 Class 1, EMD-26494 for BG24_{iGL-LC}-GT1_{N276gp120}-10-1074 Class 2, EMD-26495 for BG24_{iGL-LC}-GT1_{N276gp120}-10-1074 Class 3, and 7UGQ and EMD-26496 for BG24_{CDRL-iGL-6405}-10-1074. Local refinement maps used to model CDRL1s of BG24-derivatives have been deposited with PDB and EMDB accession codes for each respective structure.

References

1. Escolano, A. *et al.* Immunization expands B cells specific to HIV-1 V3 glycan in mice and macaques. *Nature* **570**, 468–473 (2019).
2. Lee, J. H. & Crotty, S. HIV vaccinology: 2021 update. *Semin. Immunol.* **51**, 101470 (2021).
3. Medina-Ramírez, M. *et al.* Design and crystal structure of a native-like HIV-1 envelope trimer that engages multiple broadly neutralizing antibody precursors in vivo. *J. Exp. Med.* **214**, 2573–2590 (2017).
4. Sanders, R. W. & Moore, J. P. Native-like Env trimers as a platform for HIV-1 vaccine design. *Immunol. Rev.* **275**, 161–182 (2017).
5. Zhou, T. Structural Basis for Broad and Potent Neutralization of HIV-1 by Antibody VRC01. 1–8 (2010).
6. Schommers, P. *et al.* Restriction of HIV-1 Escape by a Highly Broad and Potent Neutralizing Antibody. *Cell* **180**, 471-489.e22 (2020).
7. Huang, J. *et al.* Identification of a CD4-Binding-Site Antibody to HIV that Evolved Near-Pan Neutralization Breadth. *Immunity* **45**, 1108–1121 (2016).
8. Scheid, J. F. *et al.* Sequence and Structural Convergence of Broad and Potent HIV Antibodies That Mimic CD4 Binding. *Science* **333**, 1633–1637 (2011).
9. Wu, X. *et al.* Rational Design of Envelope Identifies Broadly Neutralizing Human Monoclonal Antibodies to HIV-1. *Science* **329**, 856–861 (2010).
10. Bar, K. J. *et al.* Effect of HIV Antibody VRC01 on Viral Rebound after Treatment Interruption. *N. Engl. J. Med.* **375**, 2037–2050 (2016).
11. Bar-On, Y. *et al.* Safety and antiviral activity of combination HIV-1 broadly neutralizing antibodies in viremic individuals. *Nat. Med.* **24**, 1701–1707 (2018).

12. Miner, M. D., Corey, L. & Montefiori, D. Broadly neutralizing monoclonal antibodies for HIV prevention. *J. Int. AIDS Soc.* **24**, (2021).
13. Mendoza, P. *et al.* Combination therapy with anti-HIV-1 antibodies maintains viral suppression. *Nature* 1–21 (2018) doi:10.1038/s41586-018-0531-2.
14. VRC 602 Study Team *et al.* Safety, pharmacokinetics and neutralization of the broadly neutralizing HIV-1 human monoclonal antibody VRC01 in healthy adults. *Clin. Exp. Immunol.* **182**, 289–301 (2015).
15. Lynch, R. M. *et al.* Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci. Transl. Med.* **7**, (2015).
16. Mayer, K. H. *et al.* Safety, pharmacokinetics, and immunological activities of multiple intravenous or subcutaneous doses of an anti-HIV monoclonal antibody, VRC01, administered to HIV-uninfected adults: Results of a phase 1 randomized trial. *PLOS Med.* **14**, e1002435 (2017).
17. Escolano, A., Dosenovic, P. & Nussenzweig, M. C. Progress toward active or passive HIV-1 vaccination. *J. Exp. Med.* **214**, 3–16 (2017).
18. West Jr, A. P. Structural basis for germ-line gene usage of a potent class of antibodies targeting the CD4-binding site of HIV-1 gp120. 1–8 (2012) doi:10.1073/pnas.1208984109/-/DCSupplemental.
19. Xiao, X. *et al.* Germline-like predecessors of broadly neutralizing antibodies lack measurable binding to HIV-1 envelope glycoproteins: Implications for evasion of immune responses and design of vaccine immunogens. *Biochem. Biophys. Res. Commun.* **390**, 404–409 (2009).

20. Bonsignori, M. *et al.* Analysis of a Clonal Lineage of HIV-1 Envelope V2/V3 Conformational Epitope-Specific Broadly Neutralizing Antibodies and Their Inferred Unmutated Common Ancestors. *J. Virol.* **85**, 9998–10009 (2011).
21. Burton, D. R. Advancing an HIV vaccine; advancing vaccinology. *Nat. Rev. Immunol.* **19**, 77–78 (2019).
22. Borst, A. Germline VRC01 antibody recognition of a modified clade C HIV-1 envelope trimer and a glycosylated HIV-1 gp120 core. 1–32 (2018).
23. Scharf, L. *et al.* Structural basis for HIV-1 gp120 recognition by a germ-line version of a broadly neutralizing antibody. *Proc. Natl. Acad. Sci.* **110**, 6049–6054 (2013).
24. Zhou, T. *et al.* Structural Repertoire of HIV-1-Neutralizing Antibodies Targeting the CD4 Supersite in 14 Donors. *Cell* **161**, 1280–1292 (2015).
25. Jardine, J. G. *et al.* HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. *science* **351**, 1458–1463 (2016).
26. Zhou, T. *et al.* Quantification of the Impact of the HIV-1-Glycan Shield on Antibody Elicitation. *Cell Rep.* **19**, 719–732 (2017).
27. Gristick, H. B. *et al.* Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site. *Nat. Publ. Group* **23**, 906–915 (2016).
28. 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).
29. Jardine, J. G. *et al.* Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen. *Science* **349**, 156–161 (2015).

30. Barnes, C. O. *et al.* A naturally arising broad and potent CD4-binding site antibody with low somatic mutation. *BioRxiv* 51 (2022)
doi:<https://www.biorxiv.org/content/10.1101/2022.03.16.484662v1>.
31. Huang, D. *et al.* B cells expressing authentic naive human VRC01-class BCRs can be recruited to germinal centers and affinity mature in multiple independent mouse models. *Proc. Natl. Acad. Sci.* **117**, 22920–22931 (2020).
32. Umotoy, J. *et al.* Rapid and Focused Maturation of a VRC01-Class HIV Broadly Neutralizing Antibody Lineage Involves Both Binding and Accommodation of the N276-Glycan. *Immunity* **51**, 141-154.e6 (2019).
33. 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**, e1003618-20 (2013).
34. Mouquet, H. *et al.* Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. *Proc. Natl. Acad. Sci.* **109**, E3268–E3277 (2012).
35. Havenar-Daughton, C. The human naive B cell repertoire contains distinct subclasses for a germline-targeting HIV-1 vaccine immunogen. 1–15 (2018).
36. Liao, H.-X. *et al.* Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* **496**, 469–476 (2013).
37. Parks, K. R. *et al.* Overcoming Steric Restrictions of VRC01 HIV-1 Neutralizing Antibodies through Immunization. *CellReports* **29**, 3060-3072.e7 (2019).
38. Stamatatos, L., Pancera, M. & McGuire, A. T. Germline-targeting immunogens. *Immunol. Rev.* **275**, 203–216 (2017).

39. 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).
40. Jardine, J. *et al.* Rational HIV Immunogen Design to Target Specific Germline B Cell Receptors. *Science* **340**, 711–716 (2013).
41. Seydoux, E. *et al.* Development of a VRC01-class germline targeting immunogen derived from anti-idiotypic antibodies. *Cell Rep.* **35**, 109084 (2021).
42. Dosenovic, P. *et al.* Anti-idiotypic antibodies elicit anti-HIV-1–specific B cell responses. *J. Exp. Med.* **216**, 2316–2330 (2019).
43. Lin, Y.-R. *et al.* HIV-1 VRC01 Germline-Targeting Immunogens Select Distinct Epitope-Specific B Cell Receptors. *Immunity* **53**, 840-851.e6 (2020).
44. Jardine, J. G. *et al.* Minimally Mutated HIV-1 Broadly Neutralizing Antibodies to Guide Reductionist Vaccine Design. *PLoS Pathog.* **12**, e1005815-33 (2016).
45. Briney, B. *et al.* Tailored Immunogens Direct Affinity Maturation toward HIV Neutralizing Antibodies. *Cell* **166**, 1459-1470.e11 (2016).
46. Yang, Z., Wang, H., Liu, A. Z., Gristick, H. B. & Bjorkman, P. J. Asymmetric opening of HIV-1 Env bound to CD4 and a coreceptor-mimicking antibody. *Nat. Struct. Mol. Biol.* **26**, 1167–1175 (2019).
47. Escolano, A. *et al.* Sequential immunization of macaques elicits heterologous neutralizing antibodies targeting the V3-glycan patch of HIV-1 Env. *Sci. Transl. Med.* **13**, eabk1533 (2021).
48. Joyce, M. G. *et al.* Soluble Prefusion Closed DS-SOSIP.664-Env Trimers of Diverse HIV-1 Strains. *CellReports* **21**, 2992–3002 (2017).

49. Cupo, A. *et al.* Optimizing the production and affinity purification of HIV-1 envelope glycoprotein SOSIP trimers from transiently transfected CHO cells. *PLOS ONE* **14**, e0215106 (2019).
50. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 133–144 (2010).
51. Winn, M. D. *et al.* Overview of the CCP 4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
52. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
53. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
54. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
55. Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 531–544 (2018).
56. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
57. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
58. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

59. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
60. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
61. Scheres, S. H. W. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
62. Tan, Y. Z. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
63. Pettersen, E. F. *et al.* UCSF Chimera?A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
64. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis: UCSF ChimeraX Visualization System. *Protein Sci.* **27**, 14–25 (2018).
65. Pettersen, E. F. *et al.* UCSF CHIMERAX : Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
66. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372**, 774–797 (2007).

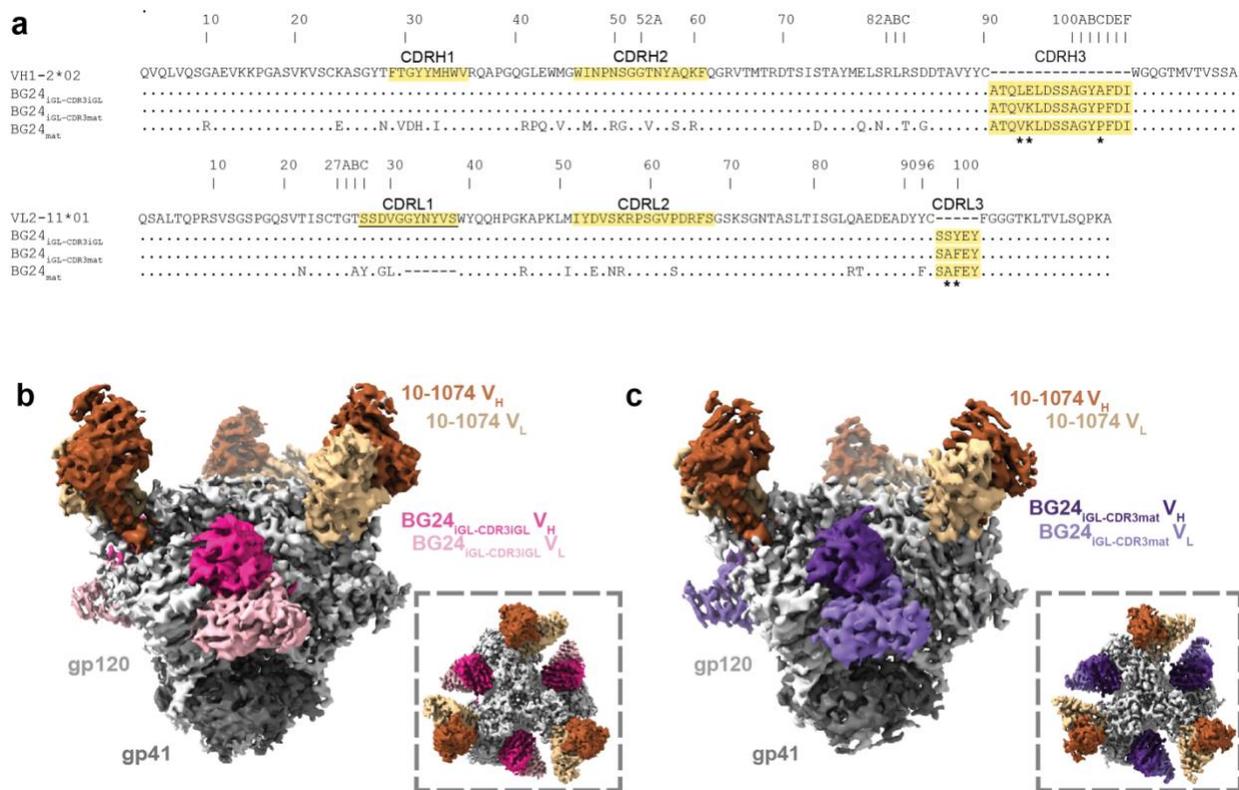


Fig. 1 | BG24_{iGL}s bind the CD4bs of the GT1 immunogen

a, Sequence alignment of V_H and V_L iGL precursors of BG24 (VH1_2*02 and VL2_11*01), BG24_{iGL-CDR3iGL}, BG24_{iGL-CDR3mat}, and BG24_{mat}. CDRs are highlighted in yellow. Asterisks (*) indicate residue differences between mature and iGL CDR3s. Underlined CDRL1 indicates sequence used for the CDRL1 in the BG24_{CDRL1-iGL} construct. **b,c**, Side and top-down (inset) views of cryo-EM density of BG24_{iGL-CDR3iGL}-GT1-10-1074 (**b**) and BG24_{iGL-CDR3mat}-GT1-10-1074 (**c**).

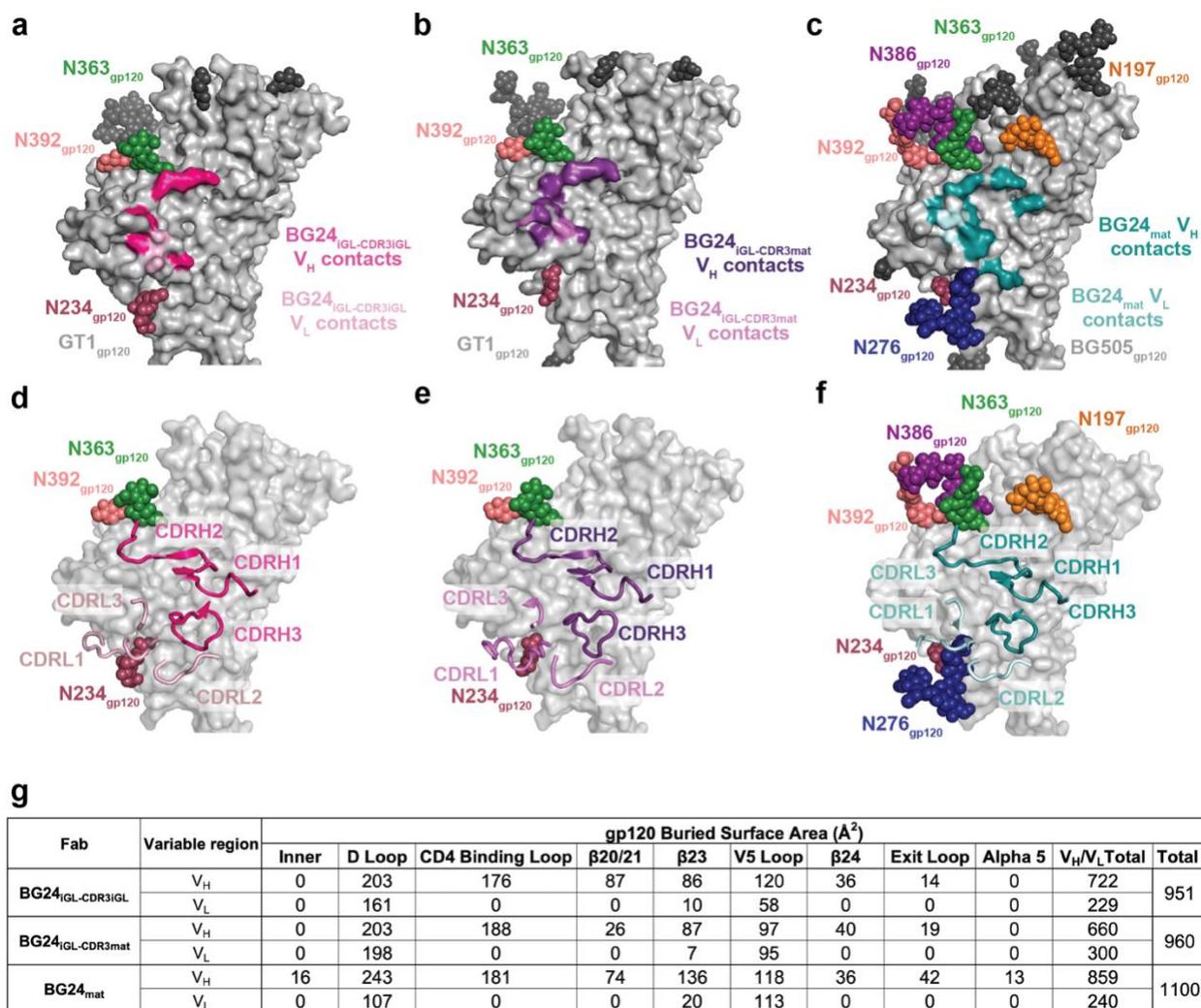


Fig. 2 | Comparison of BG24_{IGL} and BG24_{mat} CD4bs epitopes

Surface contacts made by BG24_{IGL-CDR3IGL} V_H (dark pink) and V_L (light pink) on GT1 gp120 (a), BG24_{IGL-CDR3mat} V_H (dark purple) and V_L (light purple) on GT1 gp120 (b), BG24_{mat} V_H (deep teal) and V_L (light teal) surface contacts on BG505 by BG24_{mat} (PDB 7UCF) (c). Surface representation of gp120 with cartoon representations of BG24_{IGL-CDR3mat} (d), BG24_{IGL-CDR3IGL} (e), and BG24_{mat} (f) CDR loops. g, Summary table of gp120 buried surface area (BSA) (\AA^2) calculations for BG24_{IGL-CDR3IGL}, BG24_{IGL-CDR3mat}, and BG24_{mat} at the inner domain (inner), D loop, CD4bs loop, β 20/21, β 23, V5 loop, β 24, exit loop, and alpha 5 regions of the CD4bs. BSA calculations were conducted for gp120 peptide components and did not include glycan interactions.

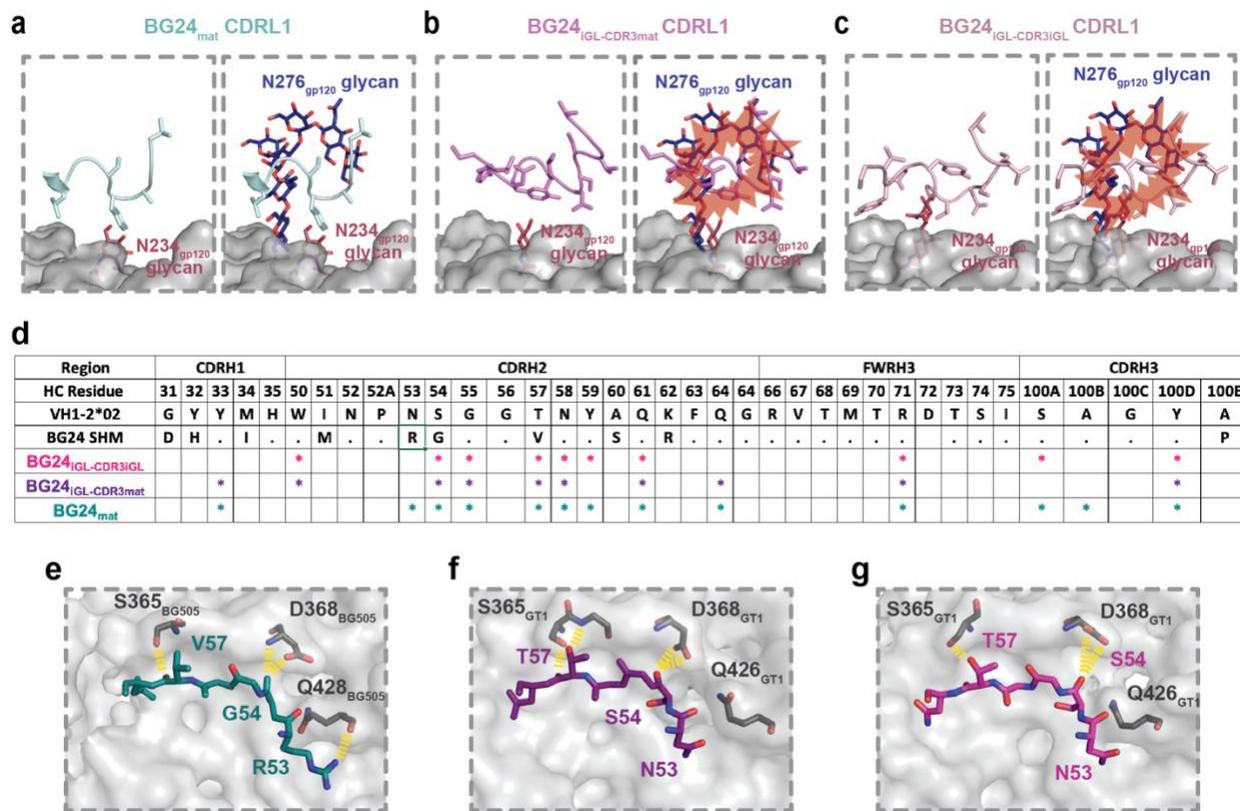


Fig. 3 | Somatic hypermutation plays a role in BG24 recognition of the CD4bs interface

gp120 surface in the vicinity of the CD4bs with cartoon representation main chain/stick side chains for the CDRL1s of **a**, BG24_{mat}, **b**, BG24_{iGL-CDR3mat}, **c**, BG24_{iGL-CDR3iGL} overlaid with the N276_{gp120} N-glycan from the BG24_{mat}-BG505 complex (PDB 7UCF). Steric clashes are represented with red bursts. **d**, Table summarizing HC paratope residues in BG24_{iGL-CDR3iGL}-GT1, BG24_{iGL-CDR3mat}-GT1 and BG24_{mat}-BG505 structures. The paratope was defined by Ab residues that make contacts with gp120 within 4 Å for each structure. Stick representations of the CDRH2 residues from **e**, BG24_{mat}, **f**, BG24_{iGL-CDR3mat}, **g**, BG24_{iGL-CDR3iGL} interacting with BG505 or GT1 gp120 residues. Yellow dashed lines indicate Ab-gp120 interactions within 4 Å.

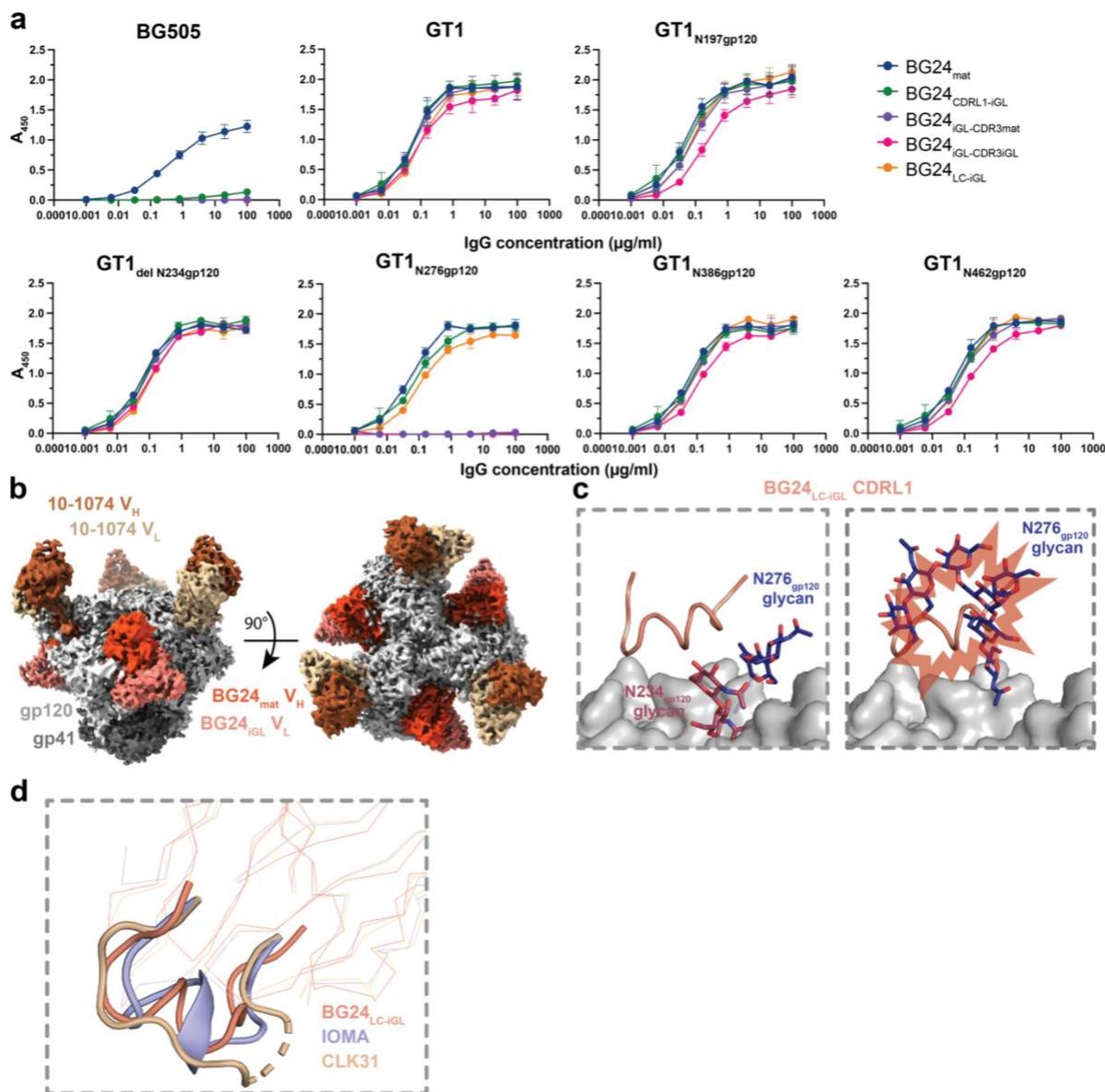


Fig. 4 | BG24_{iGL} binding is mediated by CD4bs glycans

a, ELISA to access binding of the indicated BG24 Abs to BG505, GT1, and GT1 SOSIP Envs with altered N-glycans in the CD4bs. Streptavidin plates were coated with randomly biotinylated SOSIPs and incubated with IgGs at increasing concentrations. Values are shown as mean of two individual replicates with associated error bars. **b**, Side and top-down views of cryo-EM density of BG24_{LC-iGL}-GT1-10-1074. **c**, Cartoon representation of the CDRL1 of BG24_{LC-iGL} (left) and overlaid with the N276_{gp120} N-glycan from a BG24_{mat}-BG505 (PDB 7UCF) (right). Predicted

steric clashes are indicated by red bursts. **d**, Alignment of BG24_{LC-iGL} (from the BG24_{LC-iGL}-GT1_{N276gp120}-10-1074 structure), IOMA (PDB 5T3Z), and CLK31 (PDB 6P2P) LC with CDRL1s represented in cartoon.

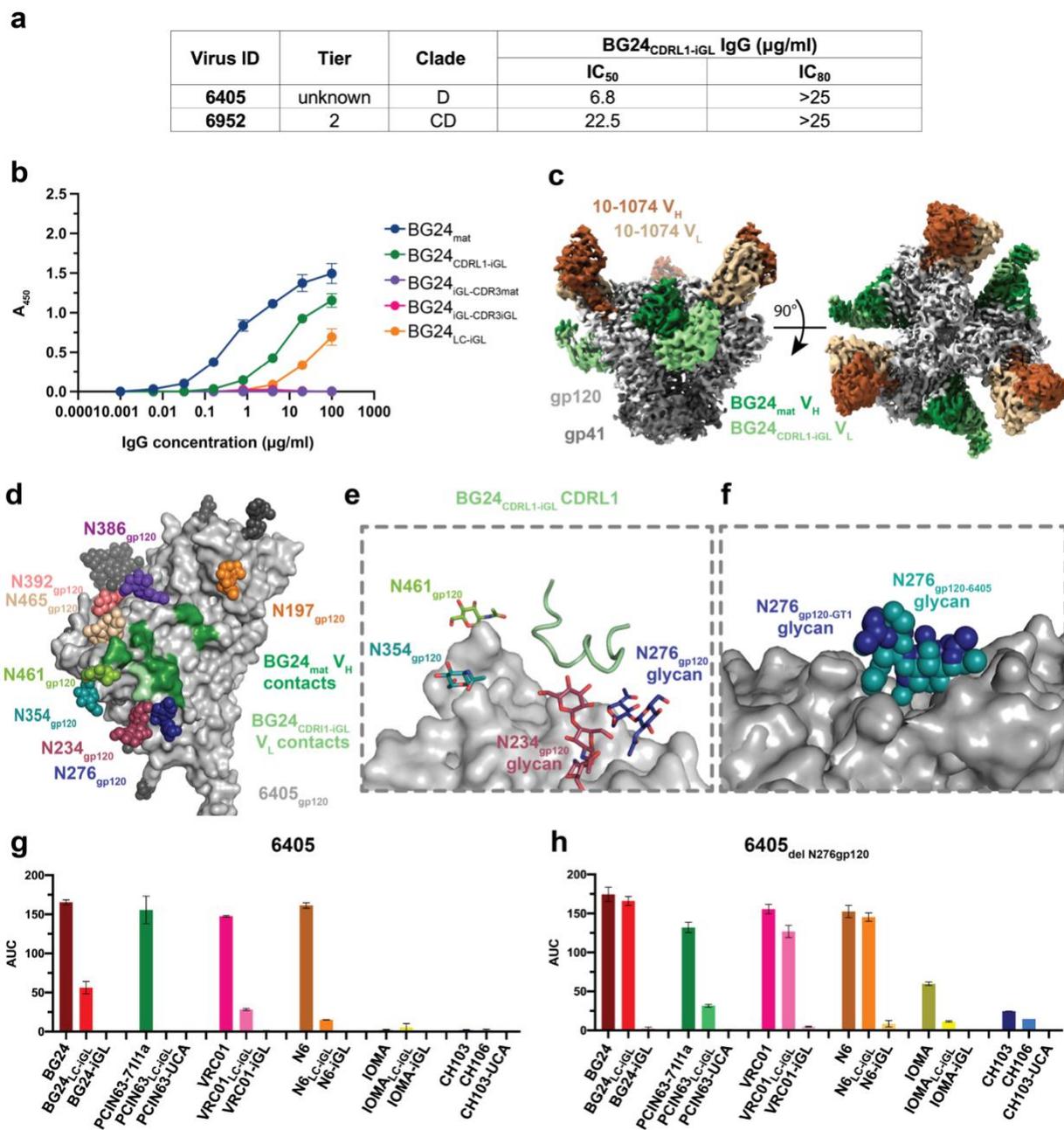
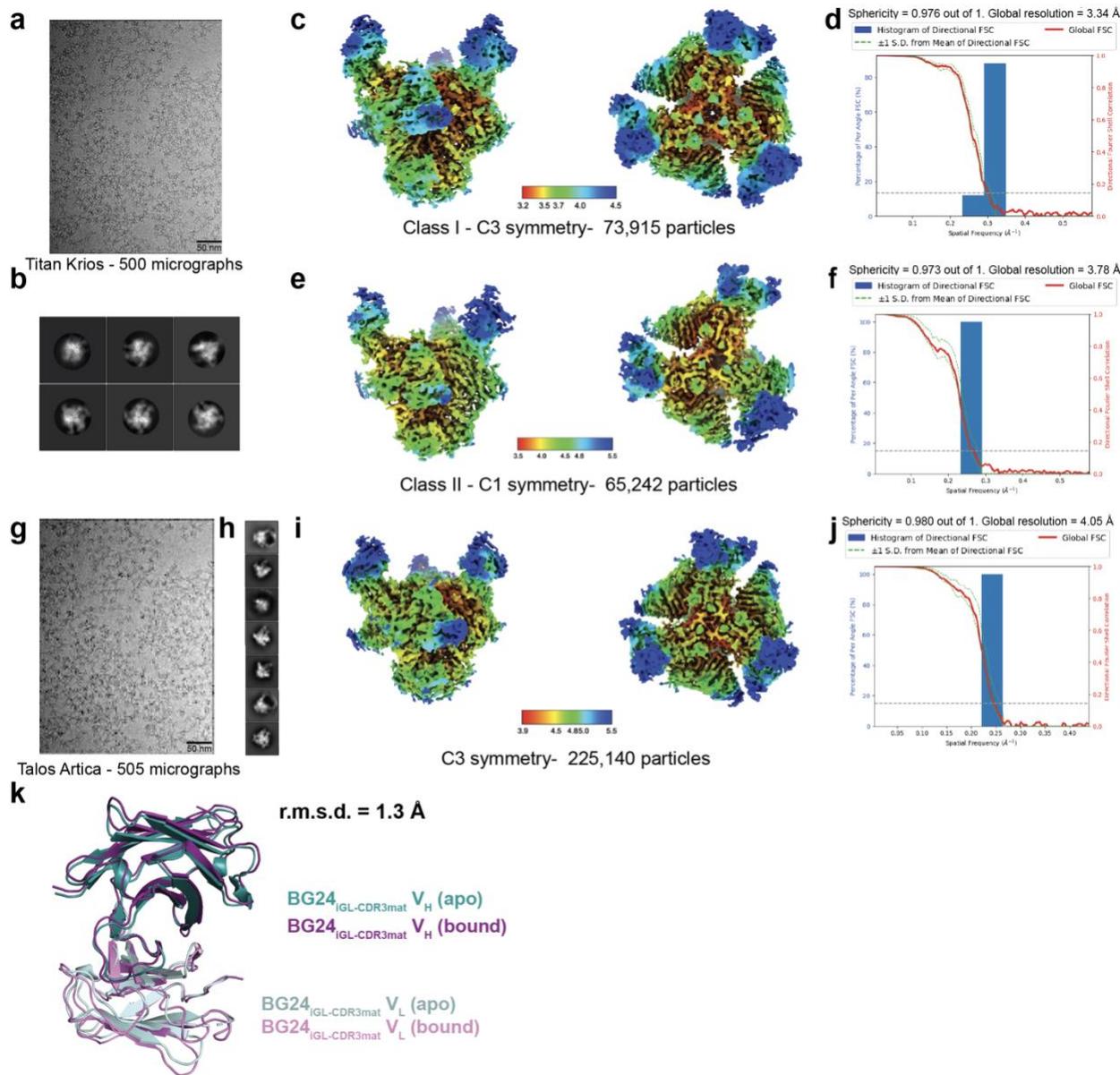


Fig. 5 | Non-engineered 6405 SOSIP recognizes BG24_{CDRL1-IgI}

a, Summary of neutralization of 6405 and 6952 pseudoviruses by BG24_{CDRL1-IgI} IgGs. **b**, ELISA to access binding of BG24-derived Abs to 6405 SOSIP. Streptavidin plates were coated with randomly biotinylated SOSIPs and incubated with BG24-derived IgGs, at increasing concentrations. **c**, Side and top-down views of cryo-EM density of BG24_{CDRL1-IgI}-6405-10-1074. **d**, Surface contacts made by BG24_{CDRL1-IgI} V_H (dark green) and V_L (light green) on 6405 gp120.

e, Cartoon representation for the CDLR1 of BG24_{CDRL1-iGL}. **f**, Alignment of GT1_{N276gp120} and 6405 gp120s in surface representation and N276 glycans in sphere representation. **g,h**, Summary for area under the curve (AUC) values derived from ELISAs that accessed binding of CD4bs IgGs to **g**, 6405 and **h**, 6405_{delN276gp120} SOSIPs. Streptavidin plates were coated with randomly biotinylated SOSIPs and incubated with CD4bs IgGs at increasing concentrations. Values are shown as mean of two individual replicates with associated error bars.

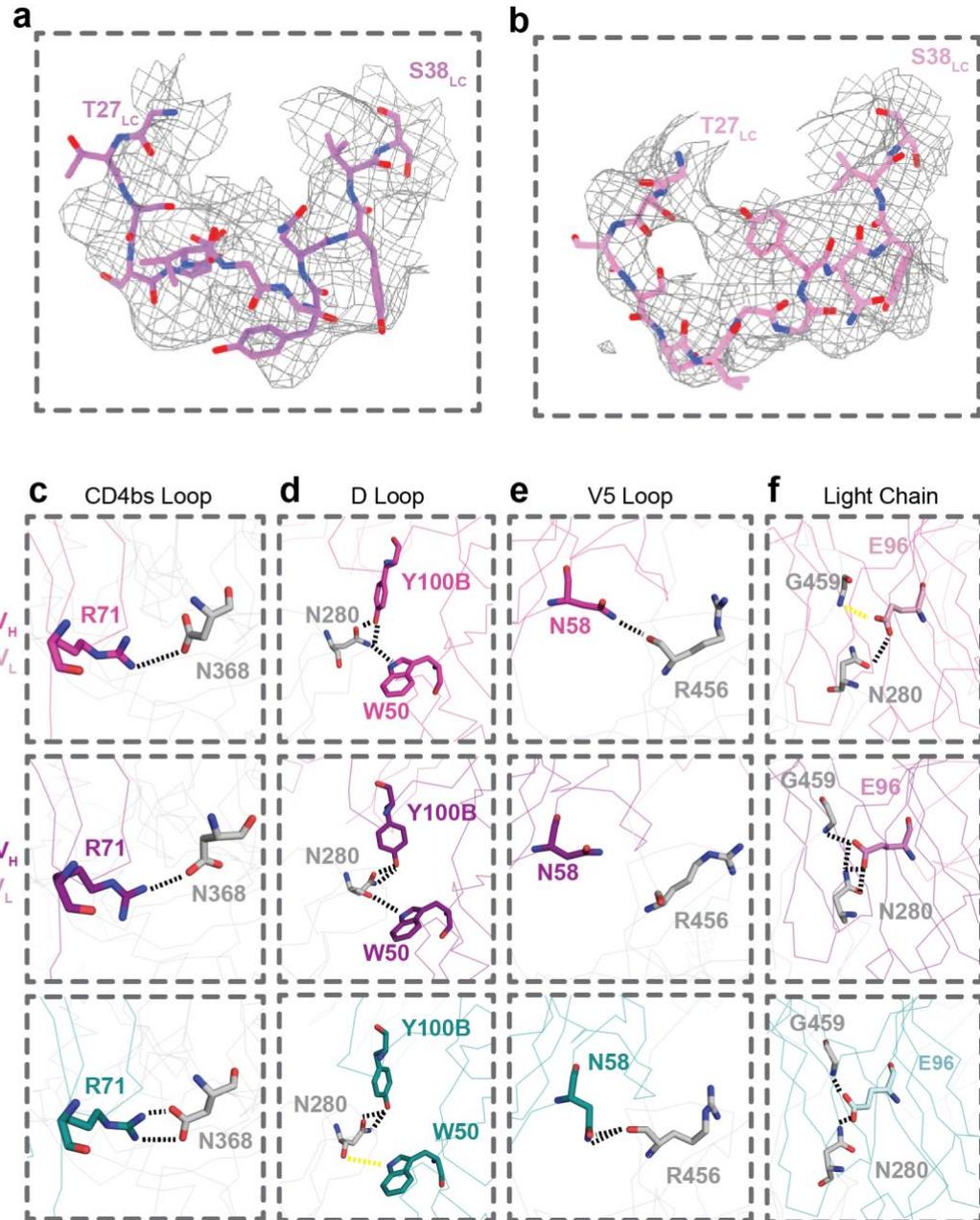


Extended Data Fig. 1 | Cryo-EM data processing and validation for BG24_{iGL}- GT1-10-1074 complexes and BG24_{iGL-CDR3mat} Fab alignment

a-j, Representative micrograph, cryo-EM 2D class averages, local resolution estimations, and gold-standard Fourier shell correlation (FSC) plots for **a-f**, BG24_{iGL-CDR3iGL}- GT1-10-1074 and **g-j**, BG24_{iGL-CDR3iGL}-GT1-10-1074. For the GT1-BG24_{iGL-CDR3iGL}-10-1074 dataset, two classes were resolved: class I with three BG24_{iGL-CDR3iGL} Fabs bound to GT1, and class II with two

BG24_{iGL-CDR3mat} Fabs bound to GT1. **k**, Alignment and r.m.s.d. (Å) of the crystal structure of apo

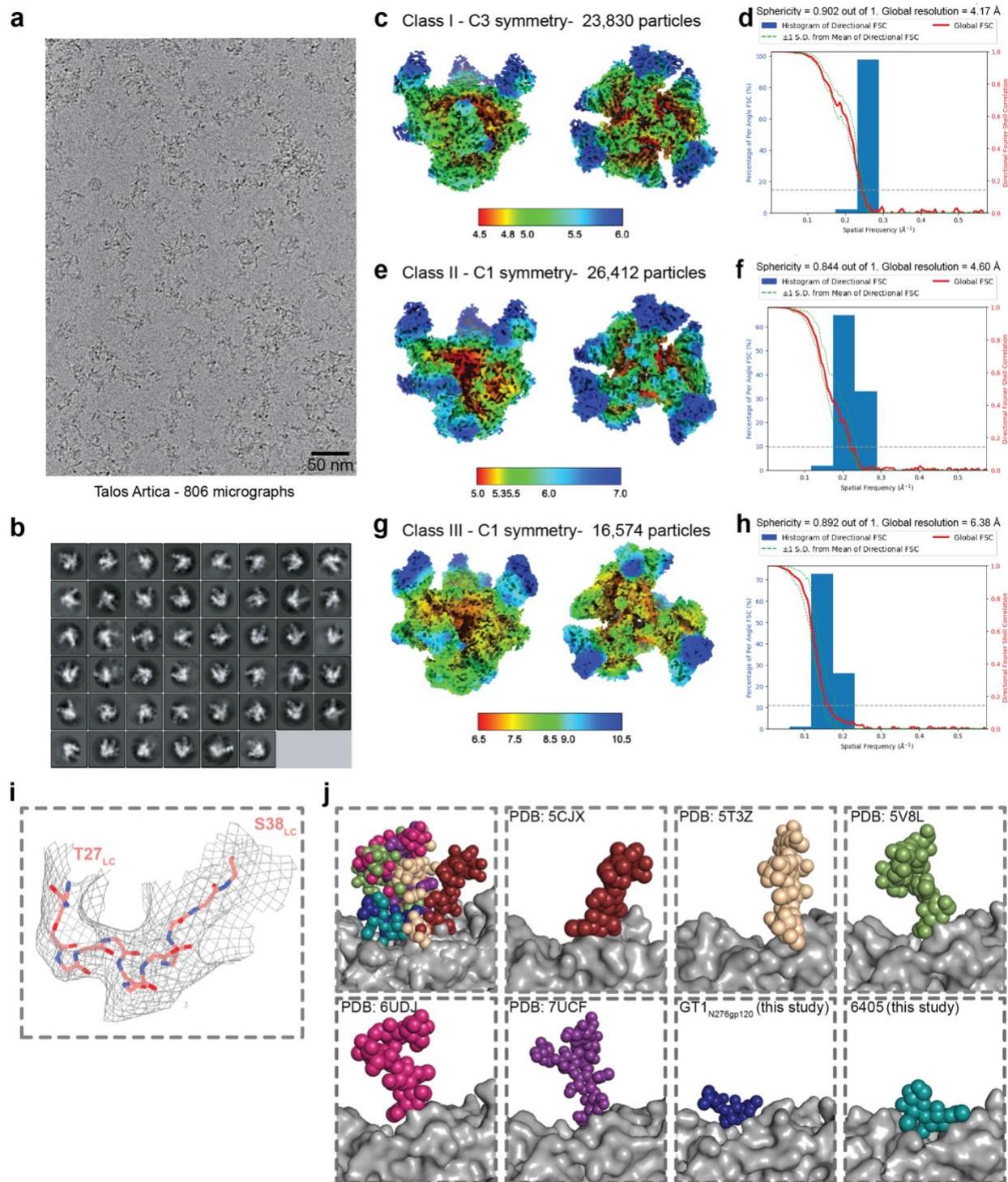
BG24_{iGL-CDR3mat} Fab and the BG24_{iGL-CDR3mat} Fab bound to GT1 in the cryo-EM structure.



Extended Data Fig. 2 | BG24_{iGL} CDR1 local density and VH1-2*02 signature contacts comparison among BG24_{iGL} and BG24_{mat} structures

Local density for CDRL1 with modeled LC residues T27-S38_{LC} (stick representation) for **a**, BG24_{iGL-CDR3mat} and **b**, BG24_{iGL-CDR3IGL} contoured to 3.8 and 3.6 σ , respectively. Interactions in **c**, CD4bs, **d**, D Loop, **e**, V5 loop, and **f**, LC for BG24_{iGL-CDR3IGL}-GT1_{gp120}, BG24_{iGL-CDR3mat}-GT1_{gp120},

and BG24_{mat}-BG505_{gp120} structures. Contacts between atoms within 4 Å are represented by black dotted lines; contacts between 4-5 Å are represented by yellow dotted lines.



Extended Data Fig. 3 | Cryo-EM data processing and validation for BG24_{LC-iGL}-GT1_{N276gp120}-10-1074 complex, BG24_{LC-iGL} CDRL1 local density, and analysis of N276_{gp120} glycan flexibility

a-h, Representative micrograph, cryo-EM 2D class averages, local resolution estimations, and gold-standard Fourier shell correlation (FSC) plots for BG24_{LC-iGL}-GT1_{N276gp120}-10-1074. For this dataset, three classes were resolved: class I with three BG24_{LC-iGL} Fabs bound to GT1_{N276gp120}, class II with two BG24_{LC-iGL} Fabs bound to GT1_{N276gp120}, and class III with one BG24_{LC-iGL} Fab bound to GT1_{N276gp120}. **i**, Local density for CDRL1 with modeled backbone for LC residues T27-S38_{LC} (stick representation) for BG24_{LC-iGL} contoured to 5.4 σ . **j**, Comparison and overlay of the N276_{gp120} glycan from existing Env structures and GT1_{N276gp120} and 6405 from this study.

```

Consensus XNLWTVYYGVPVWKAETTLFCASDAKAYETEKHNWVATHACVPTDPNPQEIHLNVTEEFNMWKNMVEQMHTDIISL
BG505     S.....
GT1      E.....K.....
6405     ND.....E.K.....S...A..I.....VE.V...N.....D..D...E.....

Consensus WDQSLKPCVKLTPLCVTLQCTNVNITDD-----MRGELKNCSFNMTTEL RDKXQKVHALFYRLDIVPINENQ
BG505     .....K...YS.....V.Q.....
GT1      E.....R.....K.....
6405     ..E.....SEMK.GTRGASNSTGTNNGTISSTDRM....A..VVN..QK.....E..GN--

Consensus -----XNTEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLL
BG505     GNRSNNS.K.....
GT1      .....S.....A.....
6405     -----NS...M.....A.S...IT.....L..R..Q.....TN..S.....

          Loop D
          276
Consensus LNGSLAEEVMIRSENITNNAKNILVQFNTPVQINCTRPNNTRKSIRIGPGQAFYATG--DIIGDIRQAHCVSKATWN
BG505     .....
GT1      .....D.R.....W.....
6405     .....DII....L.D.T.T.I..L.KSIA.....Y...QR.H.....IWTNDARV.V.....RVD..

          CD4bs Loop          V4 Loop
          352  363          386  392
Consensus ETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRI
BG505     .....
GT1      .....D.....
6405     NMTQWAATK.GSLY-.RST.I.NHA...P.I...T.....DIMNT---.SPNNTDPI...Q...

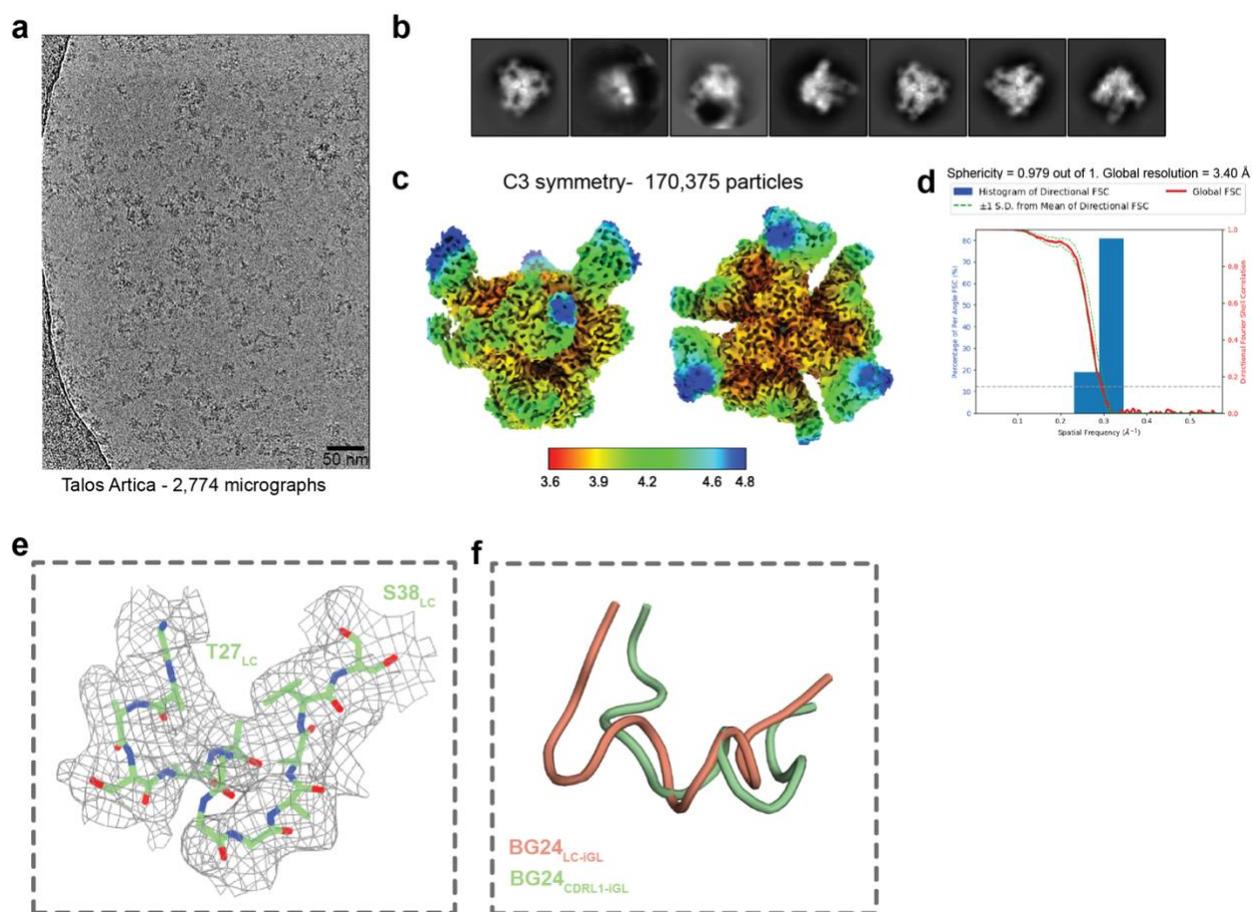
          B20/B21 Sheet          B23  V5 Loop  B24
          461  465
Consensus KQIINMWQRIGQAMYAPPIQGVIXCSNITGLILTRDGGSTNSITETFRPGGGMDRDNWRSELYKYKVVKIEPLGVAPT
BG505     .....R.V.....-.....
GT1      .....WMCQQYHR.....D.-.....S.....
6405     .....GV.R.I....A.Q.L.S.....L.....AD..SHN...N.K.....I...

Consensus RCKRRVVG
BG505     .....
GT1      .....
6405     KAR....E

```

Extended Data Fig. 4 | Sequence Alignment of BG505, GT1, and 6405 gp120

The sequence alignment for BG505, GT1, and 6405 gp120 residues and assigned consensus sequence. PNGS sequons in the consensus sequence are boxed in red, with residues numbers corresponding to Asn in the PNGS indicated above the sequence. PNGSs that are not conserved are highlighted in yellow, again with residues numbers corresponding to the Asn in the PNGS indicated above the sequence.



Extended Data Fig. 5 | Cryo-EM data processing and validation for BG24_{CDRL1-iGL}-6405-10-1074 complex and BG24_{CDRL1-iGL} CDRL1 local density

a-d, Representative micrograph, cryo-EM 2D class averages, local resolution estimations, and gold-standard Fourier shell correlation (FSC) plots for BG24_{CDRL1-iGL}-6405-10-1074. **e**, Local density for CDRL1 with modeled backbone for residues T27-S38_{LC} (stick representation) for BG24_{CDRL1-iGL} contoured to 4.4 σ . **f**, Alignment of BG24_{iGL} CDRL1s from the BG24_{LC-iGL}-GT1_{N276gp120}-10-1074 structure and from the BG24_{CDRL1-iGL}-6405-BG24_{LC-iGL}-10-1074 structure. CDRL1s are represented in cartoon.

Supplementary table 1. Cryo-EM data collection, refinement, and validation statistics

	BG24-iGL _{CDR3iGL} - GT1-10-1074	BG24-iGL _{CDR3mat} - GT1-10-1074	BG24 _{LC-iGL} - GT1 _{N276gp120} -10-1074	BG24 _{CDRL1-iGL} -6405-10-1074
EMDB:	EMD-26490	EMD-26492	EMD-26493	EMD-26496
PDB:	7UGN	7UGO	7UGP	7UGQ
Data collection and processing				
Magnification*	105,000x	73,000x	45,000x	45,000x
Voltage (kV)	300	200	200	200
Electron exposure (e ⁻ /Å ²)	60	60	60	60
Defocus range (μm)	1.2-3.0	1.2-3.0	1.2-3.0	1.2-3.0
Pixel size (Å)	0.8654	1.436	0.869	0.869
Recording mode	Counting	Counting	Counting	Counting
Symmetry imposed	C3	C3	C3	C3
Initial particle images (no.)	139,157	422,161	178,814	770,375
Final particle images (no.)	73,915	225,140	23,830	170,897
Overall map resolution (Å) (masked/unmasked)**	3.4 (3.7)	4.1 (5.0)	4.2 (4.8)	3.4 (3.7)
Refinement				
Initial model used (PDB code)	5T3Z	5T3Z	5T3Z	5T3Z
Map and model CC	0.84	0.82	0.73	0.79
Map sharpening <i>B</i> factor (Å ²)	-77	-150	-138	-119
Model composition				
Protein residues	3150	3138	3111	3162
Glycan residues	67	87	66	135
Validation				
MolProbity score	1.64	1.72	1.97	2.08
Clashscore	5.95	5.97	11.63	13.18
Poor rotamers (%)	0.04	0	0.11	0.26
Ramachandran plot				
Favored (%)	95.5	94.2	94.3	92.9
Allowed (%)	4.5	5.8	5.7	7.1
Disallowed (%)	0	0	0	0
RMS deviations				
Length (Å)	0.002	0.002	0.002	0.004
Angles (°)	0.508	0.491	0.539	0.606

* Nominal magnification; ** FSC threshold 0.143

Supplementary table 2. X-ray data collection and refinement statistics (molecular replacement)

PDB ID	iGL BG24 Fab (12-2, SSRL) 7UGM
Data collection^a	
Space group	P2 ₁ 2 ₁ 2 ₁
Unit cell (Å)	53.2, 70.8, 134.9
α, β, γ (°)	90, 90, 90
Wavelength (Å)	1.0
Resolution (Å)	38-1.4 (1.42-1.4)
Unique Reflections	101,139 (49,320)
Completeness (%)	100 (99.9)
Redundancy	19.7 (19.2)
CC _{1/2} (%)	99.3 (67.2)
$\langle I/\sigma I \rangle$	11.0 (1.2)
Mosaicity (°)	0.17
R _{merge} (%)	16 (103)
R _{pim} (%)	3.7 (237)
Wilson <i>B</i> -factor	14.2
Refinement and Validation	
Resolution (Å)	38-1.4
Number of atoms	
Protein	3,189
Water	447
R _{work} /R _{free} (%)	19.3/20.8
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.93
MolProbity score	1.42
Clashscore (all atom)	4.6
Poor rotamers (%)	0.3
Ramachandran plot	
Favored (%)	96.9
Allowed (%)	3.1
Disallowed (%)	0.2
Average <i>B</i> -factor (Å)	
Protein	24.3
Water	32.7