BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

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

Leopold Kong,[a,b,c] Alba Torrents de la Peña,[d] Marc C. Deller,[a,c,e] Fernando Garces,[a,b,c] Kwinten Sliepen,[d] Yuanzi Hua,[a] Robyn L. Stanfield,[a,b,c] Rogier W. Sanders[d] and Ian A. Wilson[a,b,c,e,f]*

[a]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California, USA, [b]Scripps Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery, [c]International AIDS Vaccine Initiative Neutralizing Antibody Center and Collaboration for AIDS Vaccine Discovery, [d]Department of Medical Microbiology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, [e]Joint Center for Structural Genomics, http://www.jcsg.org, and [f]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California, USA. *Correspondence e-mail: wilson@scripps.edu

Received 24 June 2015
Accepted 22 July 2015

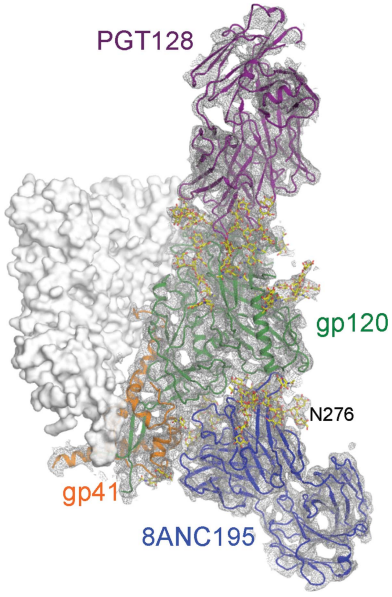
Edited by R. McKenna, University of Florida, USA

Keywords: HIV-1; SOSIP gp140; broadly neutralizing antibodies; vaccine design; glycans.

PDB reference: BG505 SOSIP gp140 HIV-1 Env trimer bound to PGT128 and 8ANC195, 5c7k

Supporting information: this article has supporting information at journals.iucr.org/d

The HIV-1 envelope gp160 glycoprotein (Env) is a trimer of gp120 and gp41 heterodimers that mediates cell entry and is the primary target of the humoral immune response. Broadly neutralizing antibodies (bNAbs) to HIV-1 have revealed multiple epitopes or sites of vulnerability, but mapping of most of these sites is incomplete owing to a paucity of structural information on the full epitope in the context of the Env trimer. Here, a crystal structure of the soluble BG505 SOSIP gp140 trimer at 4.6 Å resolution with the bNAbs 8ANC195 and PGT128 reveals additional interactions in comparison to previous antibody–gp120 structures. For 8ANC195, in addition to previously documented interactions with gp120, a substantial interface with gp41 is now elucidated that includes extensive interactions with the N637 glycan. Surprisingly, removal of the N637 glycan did not impact 8ANC195 affinity, suggesting that the antibody has evolved to accommodate this glycan without loss of binding energy. PGT128 indirectly affects the N262 glycan by a domino effect, in which PGT128 binds to the N301 glycan, which in turn interacts with and repositions the N262 glycan, thereby illustrating the important role of neighboring glycans on epitope conformation and stability. Comparisons with other Env trimer and gp120 structures support an induced conformation for glycan N262, suggesting that the glycan shield is allosterically modified upon PGT128 binding. These complete epitopes of two broadly neutralizing antibodies on the Env trimer can now be exploited for HIV-1 vaccine design.



1. Introduction

Despite decades of intensive research, no effective vaccine has been developed for *Human immunodeficiency virus 1* (HIV-1). A major challenge for vaccine design is the hypervariability of this retrovirus, which leads to rapid immune escape, in concert with the large numbers of viral subtypes (clades) and strains that must be taken into account. The HIV-1 envelope glycoprotein gp160 (Env) is the only viral protein exposed on the virion surface and is therefore the main target of the neutralizing antibody response. Env is especially diverse and contains hypervariable loops that are also prone to insertions and deletions (Korber *et al.*, 2001). Env is a trimer of heterodimers consisting of gp120 and gp41 glycoproteins that mediate receptor binding and membrane fusion, respectively. Env is also extensively covered by a dense canopy of N-linked

© 2015 International Union of Crystallography

glycans that are thought to shield the underlying protein from antibody recognition (Wei *et al.*, 2003; Wyatt *et al.*, 1998).

Until recently, structural characterization of Env has been challenging, often requiring extensive loop truncations to derive monomeric gp120 core constructs that are amenable to crystallization (Kwong *et al.*, 1998, 1999). Despite their limitations, structures of the gp120 core bound to antibodies have inspired a rational design approach to optimize the presentation of structurally defined antibody epitopes on vaccine immunogens (Burton *et al.*, 2012; Jardine *et al.*, 2013). This approach has been further driven by the discovery and isolation of exceptionally potent, broadly neutralizing antibodies (bNAbs; Walker *et al.*, 2009, 2011; Scheid *et al.*, 2011), some of whose epitopes overlap on several key supersites of vulnerability (Kong *et al.*, 2013; Zhou *et al.*, 2010; McLellan *et al.*, 2011; Scharf *et al.*, 2014).

One major site of vulnerability is situated at the base of the V3 loop of gp120, centered on the N-linked glycan at Asn332 (N332; Walker *et al.*, 2011). Multiple bNAb lineages have been identified that interact with the N332 glycan, including the PGT121, PGT128 and PGT135 families of antibodies. Of particular note are the crystal structures of PGT128 (Pejchal *et al.*, 2011) and PGT135 (Kong *et al.*, 2013) bound to the gp120 outer domain fragment and gp120 core, respectively, which have provided initial templates for vaccine design. More recently, crystal structures of the bNAb PGT122, a member of the PGT121 family, bound to a stabilized gp140 trimer termed SOSIP (Julien *et al.*, 2013; Sanders *et al.*, 2013; Pancera *et al.*, 2013) have revealed a larger epitope that extends beyond the residues present in the gp120 core construct. This N332 epitope includes the V1V2 region of gp120 and encompasses N-linked glycans beyond those observed in the initial gp120–antibody structures. Recent studies investigating the impact of these glycans on the PGT121 family have revealed unexpected complexities (Garces *et al.*, 2014). While the N332 glycan plays a positive role in enhancing antibody affinity for some members of the PGT121 family, other glycans, such as N137 from the V1V2 region, can be inhibitory. For example, the bNAb PGT124 has evolved a structure and consequent binding mode that avoids or minimizes interaction with the N137 glycan, while the bNAb PGT122 has matured in a slightly different way to accommodate this glycan (Garces *et al.*, 2014). Functional studies and modeling suggest that the N137 glycan and the neighboring N156 glycan may interact with or are accommodated by PGT128, but this notion has not been demonstrated structurally (Doores *et al.*, 2015; Julien *et al.*, 2013).

Recently, multiple bNAbs have been isolated that recognize a site of vulnerability spanning the gp41 and gp120 interface (Scharf *et al.*, 2014; Huang *et al.*, 2014; Blattner *et al.*, 2014). Binding to this interface presents a challenge for bNAb recognition, as it contains gp41 glycans that are exceptionally heterogeneous relative to the glycans on gp120 (Pritchard *et al.*, 2015). The potent bNAbs 35022 and PGT151 interact with both gp41 and gp120, and their interactions with the Env trimer have been structurally characterized by crystallography (Pancera *et al.*, 2014) and electron microscopy (EM; Blattner *et al.*, 2014), respectively. The epitope for another bNAb targeting this region, 8ANC195, has been shown by negative-stain EM to extend into the gp41 region (Scharf *et al.*, 2014), but only the atomic-level details of its interaction with monomeric gp120 have been defined *via* a crystal structure (Scharf *et al.*, 2014).

In order to elucidate the full epitopes of the bNAbs PGT128 and 8ANC195 in the context of an intact Env trimer, a crystal structure of these two antibodies bound to soluble BG505 SOSIP gp140 trimer has been determined at 4.6 Å resolution. For PGT128, an indirect interaction with a highly conserved N-linked glycan is now observed, demonstrating how antibody binding may locally restructure the Env glycan shield. For 8ANC195, the binding surface on gp41, which includes extensive interactions with glycan N637, is also defined. These results extend the templates for HIV-1 immunogen design.

2. Materials and methods

2.1. Protein production and purification

The V_H and V_L chains constituting the antibody binding fragments (Fabs) of 8ANC195 and PGT128 were cloned into the expression vector phCMV3 for high-level protein expression by transient transfection of mammalian FreeStyle 293F cells. The BG505 SOSIP gp140 trimer was also cloned into the phCMV3 vector, but was produced by transient transfection of GnTI$^{-/-}$ 293S cells to ensure homogeneous oligomannose N-linked glycosylation. After 7 d, the supernatants of the PGT128- and 8ANC195-transfected cells were harvested and purified using a Lambda or a Kappa Capture Select column (BAC BV), respectively. The antibodies were further purified by ion-exchange chromatography using a 10/100 GL Mono S column (GE), followed by size-exclusion chromatography (SEC) with a HiLoad 16/600 Superdex 200 column (GE). For the SOSIP trimer protein, supernatants were also collected after 7 d of transfection and purified by 2g12 antibody affinity chromatography and SEC as described previously (Sanders *et al.*, 2013).

The SOSIP trimer was mixed with the PGT128 and 8ANC195 Fabs in a molar ratio of 1.0:1.2:1.2 (SOSIP:PGT128:8ANC195) at room temperature for 20 min to form a complex. This mixture was then deglycosylated (Julien *et al.*, 2013) with endoglycosidase H (EndoH) in 200 m*M* NaCl, 50 m*M* sodium citrate pH 5.5 at 37°C for 1 h following the manufacturer's protocol (New England Biolabs) before final purification by SEC. This limited deglycosylation step removes glycans that are not protected from interaction with the Fabs and results in cleavage between the first and second GlcNAc moieties next to the Asn.

2.2. Protein crystallization and data collection

The purified complex consisting of PGT128 and 8ANC195 bound to EndoH-treated BG505 SOSIP gp140 trimer was prepared for crystallization by buffer exchange into 50 m*M* NaCl, 20 m*M* Tris–HCl pH 7.2. The complex was concentrated to 4 mg ml^{-1} and passed through a 0.22 µm filter before being

Table 1
Data-collection and refinement statistics for the complex of BG505 SOSIP.664 gp140 with the Fabs 8ANC195 and PGT128.

Values in parentheses are for the highest resolution shell.

Data collection	
Beamline	23-ID-B, APS
Wavelength (Å)	1.0332
Space group	*I*23
Unit-cell parameters (Å)	$a = b = c = 261.1$
Resolution (Å)	50.00–4.60 (4.76–4.60)
No. of observations	74655
No. of unique reflections	16539 (1616)
Multiplicity	4.5 (4.6)
Completeness (%)	99.8 (100.0)
$\langle I\rangle/\langle\sigma(I)\rangle$	7.8 (1.7)
$R_{\rm merge}$†	0.21 (1.16)
$R_{\rm p.im.}$‡	0.12 (0.60)
$R_{\rm meas}$§	0.23 (1.31)
Refinement statistics	
Resolution (Å)	47.7–4.6 (4.73–4.60)
No. of reflections (working set)	14887 (1232)
No. of reflections (test set)	1635 (133)
$R_{\rm cryst}$¶ (%)	28.9 (34.9)
$R_{\rm free}$†† (%)	29.5 (35.2)
No. of atoms	
PGT128 (chains *A* and *B*)	3530
8ANC195 (chains *E* and *F*)	3312
gp120 (chain *C*)	3544
gp41 (chain *D*)	978
N-Linked glycan	723
Sulfate	5
Average *B* value (Å^2)	
PGT128 (chains *A* and *C*)	219
8ANC195 (chains *H* and *L*)	162
gp120 (chain *G*)	167
gp41 (chain *B*)	145
N-Linked glycan	182
Sulfate	171
Wilson *B* value (Å^2)	160
R.m.s.d. from ideal geometry	
Bond lengths (Å)	0.004
Bond angles (°)	0.97
MolProbity analysis‡‡	
All-atom clash score	9.6 [97th percentile]§§
Protein-geometry score	1.8 [100th percentile]§§
Poor rotamers (%)	0.2
Ramachandran plot (%)	
Favored	95.2
Allowed	4.7
Outliers	0.1
PDB code	5c7k

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the scaled intensity of the *i*th measurement of reflection *hkl* and $\langle I(hkl)\rangle$ is the average intensity for that reflection. ‡ $R_{\rm p.i.m.}$ is a redundancy-independent measure of the quality of intensity measurements. $R_{\rm p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the scaled intensity of the *i*th measurement of reflection *hkl*, $\langle I(hkl)\rangle$ is the average intensity for that reflection and $N(hkl)$ is the redundancy. § $R_{\rm meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the scaled intensity of the *i*th measurement of reflection *hkl*, $\langle I(hkl)\rangle$ is the average intensity for that reflection and $N(hkl)$ is the redundancy. ¶ $R_{\rm cryst} = \sum_{hkl}||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl}|F_{\rm obs}| \times 100$. †† $R_{\rm free}$ was calculated as for $R_{\rm cryst}$ but on a test set comprising 9.9% of the data, which were excluded from refinement. ‡‡ Calculated using *MolProbity* (http://molprobity.biochem.duke.edu/; Chen *et al.*, 2010). §§ Compared with structures of similar resolution using *MolProbity*.

screened for crystallization using the IAVI/JCSG/TSRI CrystalMation robot (Rigaku) at the JCSG (Elsliger *et al.*, 2010). Multiple crystallization conditions with PEG 200, PEG 400, PEG 6000 or PEG 8000 yielded crystals, but most crystals did not diffract beyond 10 Å resolution. However, crystals grown at 25°C in 0.05 *M* lithium sulfate, 0.05 *M* sodium sulfate, 20%(*w*/*v*) PEG 400, 0.05 *M* Tris–HCl pH 8.7 (JCSG Core Suite I condition 3, Qiagen) diffracted to 7 Å resolution. These crystals were further optimized using fine screening, additives, temperature variation and micro-batch seeding techniques implemented with an Oryx8 Crystallization robot (Douglas Instruments). The best diffracting crystal, from over 200 that were screened, was grown over a period of 3 d at 4°C in a crystallization reagent consisting of 0.05 *M* lithium sulfate, 0.05 *M* sodium sulfate, 26%(*w*/*v*) PEG 400, 0.1 *M* Tris–HCl pH 8.3. All crystals for X-ray data collection were cryoprotected by brief immersion in mother liquor supplemented with 40% PEG 400 prior to flash-cooling in liquid nitrogen. Diffraction data were collected to 4.6 Å resolution on beamline 23-ID-B at the Advanced Photon Source, processed with *HKL*-2000 (Otwinowski & Minor, 1997) and indexed in space group *I*23 with 100% completeness and an $\langle I\rangle/\langle\sigma(I)\rangle$ of 1.7 in the highest resolution shell. The asymmetric unit contains one gp140 protomer of the trimer bound to one 8ANC195 Fab and one PGT128 Fab, with a Matthews coefficient ($V_{\rm M}$) of 4.4 Å^3 Da^{-1} and an estimated solvent content of 72.2%. Data-collection and processing statistics are summarized in Table 1.

2.3. Structure determination and refinement

The structure of PGT128 and 8ANC195 bound to the BG5050 SOSIP trimer was solved by the molecular-replacement method using *Phaser* (McCoy *et al.*, 2007). Five search models were used: a SOSIP gp140 molecule (PDB entry 4tvp, resolution 3.5 Å; Pancera *et al.*, 2014) and the separate variable and constant regions of PGT128 (PDB entry 3tyg, resolution 3.25 Å; Pejchal *et al.*, 2011) and 8ANC195 (PDB entry 4p9h, resolution 3.0 Å; Scharf *et al.*, 2014). Refinement consisted of alternating rounds of manual model building using *Coot* v.0.7 (Emsley *et al.*, 2010) and automated refinement as implemented in *PHENIX* (Adams *et al.*, 2010).

Given the limited resolution of the data set, grouped *B*-factor refinement for each residue was used. Furthermore, positional coordinate refinement was enforced using a reference model set of restraints. The starting model for each automated refinement session in *PHENIX* was defined as the reference model for that session. Substantial modifications to the model were only made during rounds of manual model building in *Coot*. This strategy allowed the preservation of energetically preferred side-chain rotamers and Ramachandran torsion angles for most of the model, despite the low-resolution data. As defined by *MolProbity* (Chen *et al.*, 2010), only two residues were Ramachandran outliers, Val51 and Pro109 on the PGT128 light chain. Residue 51 in antibody light-chain structures is commonly observed to be a Ramachandran outlier, as it is found in a γ-turn (Stanfield *et al.*, 2003). Residues 1, 128–133 and 228–231 from the PGT128 heavy chain (chain *A*), 1–3 and 212–215 from the PGT128 light chain (chain *B*), 136–140, 186–187, 402–410 and 505–507 from gp120 (chain *C*), 512–521 and 544–563 from gp41 (chain *D*), 129–133, 187–190 and 215–219 from the 8ANC195 heavy chain (chain *E*) and 212–214 from the 8ANC195 light chain (chain *F*) are omitted owing to poor or no electron density. The final

R_{cryst} and R_{free} values converged at 28.9 and 29.5%, respectively. See Table 1 for the final refinement statistics.

The refinement of glycans at low resolution posed a special challenge owing to software limitations that prevented the preservation of the preferred chair conformation and glycosidic torsion angles, despite imposing strict restraints. Therefore, as discussed above, grouped *B*-factor refinement was implemented and model building was achieved by manual manipulation in *Coot*. As described previously (Kong *et al.*, 2013), energy-minimized rotamers of oligomannose N-linked glycan models were obtained from the GLYCAMM Programmers' Libraries (http://glycam.ccrc.uga.edu/glylib) and used as initial models. Further glycan model building was achieved by superposing energy-minimized glycans onto the glycan moieties of the initial model. Using this approach, additional glycan electron density could be more accurately modeled. Modifications to the torsion angles and geometry were made manually and were strictly monitored with the *PDB CArbohydrate REsidue* server (*PDBCARE*; http://www.glycosciences.de; Lütteke & von der Lieth, 2004).

The data set was initially processed to 4.5 Å resolution in space group *I*23, and possible twinning was indicated, with $\langle E^2 - 1 \rangle$ for noncentric reflections equal to 0.659. Merohedral twinning is possible for this space group; however, *phenix.xtriage* reported *L*-test statistics of $\langle |L| \rangle = 0.44$, $\langle L^2 \rangle = 0.27$ and an estimated twin fraction of only 0.049 for the merohedral twin law $-l, -k, -h$, so that merohedral twinning seemed unlikely. Intensity-based twinning statistics varied depending on how negative *I* values were treated during conversion to *F* values and on the resolution cutoff (10% of intensities were negative at 4.5 Å). Thus, to determine whether any twinning was present, the data were reprocessed in a lower symmetry space group, *R*3, for which pseudo-merohedral twinning in the apparent space group *I*23 is

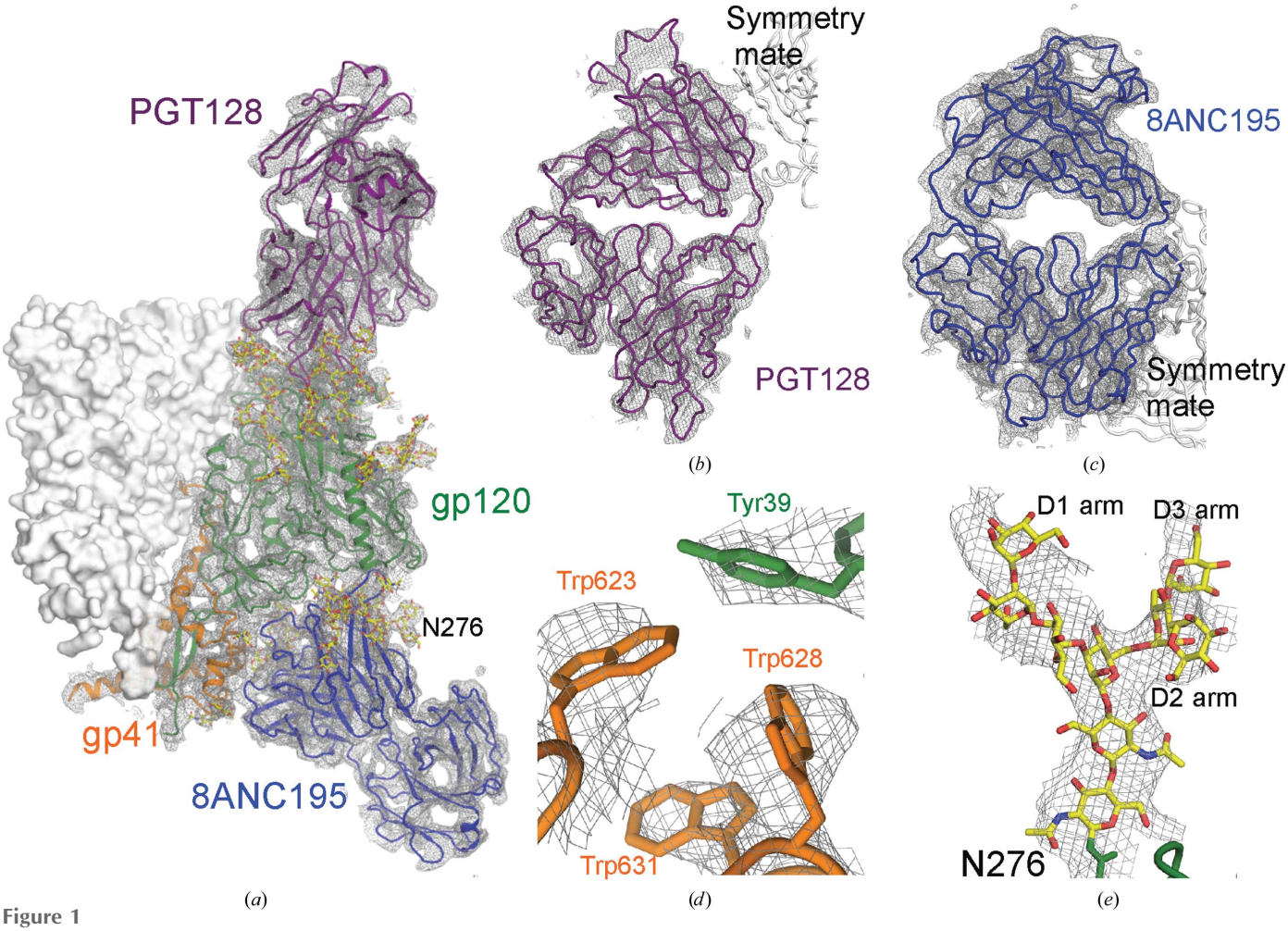


Figure 1
The overall structure and electron-density map of the Fabs PGT128 and 8ANC195 in complex with the BG505 SOSIP gp140 Env trimer. (*a*) Ribbon diagram of the asymmetric unit containing PGT128 (purple), 8ANC195 (blue), gp120 (green) and gp41 (orange). *N*-Linked glycans are shown in ball-and-stick representation with yellow C atoms, red O atoms and blue N atoms. The $2F_o - F_c$ electron-density map contoured at 1σ is shown as a gray mesh. The other two gp140 protomers that form the trimer are shown as a whitish gray molecular surface. (*b*) Enlargement of the ribbon diagram of PGT128 and electron-density map as shown in (*a*). Crystal contact with the constant region of a neighboring PGT128 symmetry mate is shown. (*c*) Enlargement of the ribbon diagram of 8ANC195 and electron-density map as shown in (*a*). The crystal contact between the variable region and a neighboring 8ANC195 symmetry mate is shown. (*d*) An electron-density map for bulky hydrophobic residues at the gp120–gp41 interface is shown. Side chains are displayed in ball-and-stick representation and colored as in (*a*). (*e*) A composite OMIT map of the N276 glycan with the final refined glycan model superimposed. The glycan model is represented as in (*a*).

possible. However, refinement with appropriate twin operators did not yield improved R_{free} values, suggesting that there was no significant twinning in this crystal. Therefore, the *I*23 space group was selected for the final data set. In addition, a slightly lower resolution cutoff (4.6 Å) was used, which improved the intensity-based statistics ($\langle E^2 - 1 \rangle = 0.681$) and resulted in a 2% lower R_{free}.

Buried molecular-surface areas were analyzed with the *Molecular Surface Package* (http://www.csb.yale.edu/userguides/graphics/msp/msu_local.html; Connolly, 1993) using a 1.7 Å probe radius and standard van der Waals radii (Gelin & Karplus, 1979). Fab residues were numbered according to the Kabat nomenclature (Martin, 1996) and gp140 was numbered using the standard HXBc2 convention.

The quality of the crystal structure was analyzed using the *wwPDB X-ray Structure Validation* pipeline (http://wwpdb-validation.wwpdb.org/validservice/). Atomic coordinates and experimental structure factors have been deposited in the PDB under accession code 5c7k.

2.4. Site-directed mutagenesis

Variants of the BG505 SOSIP.664 gp140 trimer bearing a D7324 antibody epitope sequence (GSAPTKAKRRVVQREKR) at the C-terminus of amino-acid residue 664 were generated using the QuikChange site-directed mutagenesis kit (Agilent Technologies) with PCR primers encoding the specific mutations. The mutated sequences were confirmed by DNA sequencing prior to use.

2.5. ELISA antibody binding assay

Recombinant gp140 and gp140 variants containing the D7324 epitope tag were produced by transient transfection of HEK293T cells. Pure cell supernatants were immobilized onto antibody D7324-coated half-well ELISA plates (Greiner) (50 µl per well). Antibody binding was determined as described previously (Sanders *et al.*, 2013) using a peroxidase-conjugated goat anti-human IgG antibody as the secondary antibody. The experiments were performed in duplicate.

3. Results and discussion

3.1. Overall structure

Although substantially more stable than other gp140 constructs (Sanders *et al.*, 2013), the SOSIP trimer is heavily glycosylated and contains flexible loops that make it difficult to obtain crystals that diffract to even moderate resolution. One strategy for crystallization is to screen the SOSIP trimer with a range of different antibodies to find the best diffracting crystals (Julien *et al.* 2013). The combination of a gp41–gp120 interface-targeting antibody and an N332 glycan-targeting antibody has produced the best diffracting crystals so far, to around 3.5 Å resolution (Pancera *et al.*, 2014). However, high-resolution diffraction (<3 Å) has remained elusive. Here, the gp41–gp120 interface-targeting bNAb 8ANC195 and the N332 glycan-targeting bNAb PGT128 were co-crystallized with SOSIP and the structure was solved at 4.6 Å resolution. Despite the low resolution, most of the protein C^{α} trace is visible (Fig. 1*a*). The CDR loops and elbow regions of the antibodies are also well defined by the density (Figs. 1*b* and 1*c*), with undefined portions largely being isolated to flexible loops in the constant region. Some side-chain density is clearly visible for bulky aromatic residues near interfaces (Fig. 1*d*). Fortunately, higher resolution models of gp140, PGT128 and 8ANC195 are available for use in molecular replacement and allowed higher quality refinement using these reference models as restraints (Table 1). Although the coordinates of the N-linked glycans were not included during the molecular-replacement search, density was visible directly after molecular replacement and was used to assist in *de novo* glycan

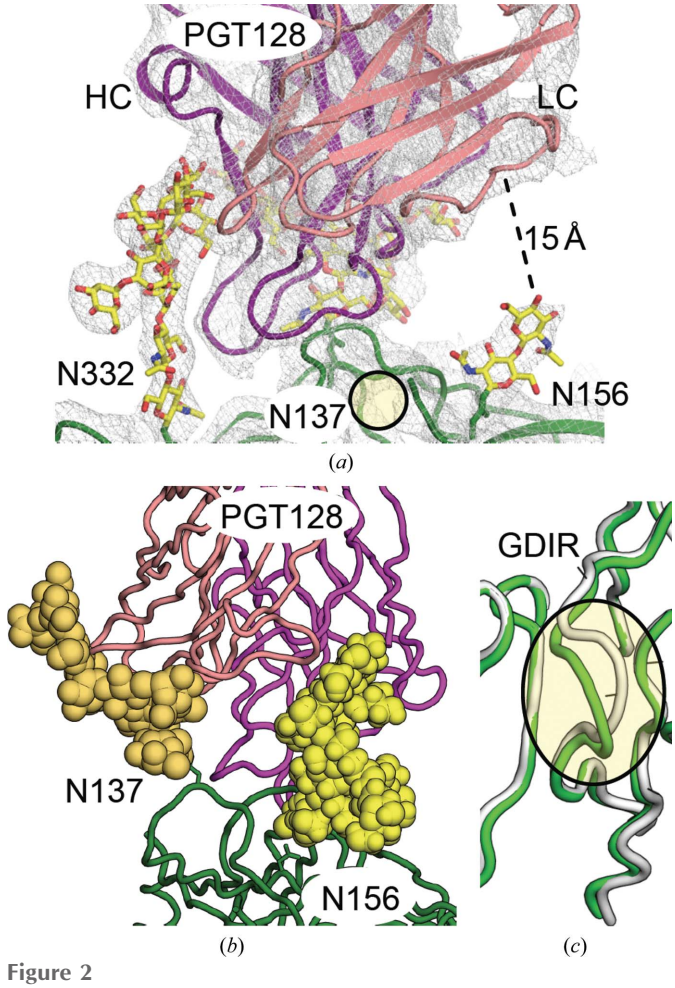


Figure 2
The PGT128 epitope on the SOSIP Env trimer. (*a*) Ribbon diagrams of the PGT128 variable region bound to the gp120 component of the Env trimer are shown with the corresponding electron density as in Fig. 1. Only the position of the N137 glycan is highlighted owing to its absence in the structure. (*b*) Glycans at positions N137 (orange spheres; PDB entry 4tvp) and N156 (yellow spheres; PDB entry 4dqo) from other crystal structures were modeled in the PGT128-bound SOSIP structure to assess potential interactions with the antibody. The view presented is rotated from the orientation in (*a*) towards the PGT128 light chain in order to display both of the modeled glycans so that they do not obscure each other. (*c*) Ribbon diagrams of the V3 region of the PGT128-bound SOSIP structure (green) and the PGT122-bound SOSIP structure (white) are superposed. Conformational variation at the 'GDIR' motif is highlighted.

model building. Glycan model building was appreciably simplified by the production of the SOSIP trimer in 293S cells, which limited the glycoforms to oligomannose types. The density of N-linked glycans at interfaces is well defined (Fig. 1*e*), allowing assignment of the D1–D3 terminal glycan arms (Vliegenthart *et al.*, 1983) when visible. The greater definition of N-linked glycans was useful in characterizing the epitopes of the glycan-dependent PGT128 and 8ANC195 antibodies.

3.2. The PGT128 epitope

In the crystal structure, PGT128 engages the SOSIP trimer in a similar fashion to the previously described interaction with the gp120 outer domain fragment (Pejchal *et al.*, 2011). In both structures, PGT128 primarily contacts the 'GDIR' protein motif at gp120 residues 324–327 and the glycans at N332 and N301. In the SOSIP structure, two additional mannose moieties are visible that were not observed in the original outer domain structure, completing the D1 arm of the N301 glycan. However, the additional mannoses do not contact PGT128.

One motivation for this study was to reveal whether PGT128 makes additional contacts with the V1V2 loop glycans at positions N137 and N156 since these interactions have been observed for PGT122, which binds at a similar position (Julien *et al.*, 2013). Previous studies showed that the PGT128 affinity is not significantly affected by mutations removing these glycans, but this result does not exclude the possibility of direct glycan contact with the antibody (Julien *et al.*, 2013; Doores *et al.*, 2015). Furthermore, the light-chain variable region of PGT128 facing these glycans is recessed relative to the extended heavy chain, making glycan accommodation possible (Julien *et al.*, 2013). However, no density is observed for the N137 glycan, and only two GlcNAc moieties are visible for the N156 glycan, the second GlcNAc of which is nearly 15 Å from the antibody (Fig. 2*a*). This N137 glycan is either disordered or possibly cleaved during the Endo H treatment, suggesting that the glycan was not protected by PGT128 binding. However, when glycans at these positions from previously published structures are superimposed onto our structure (the N137 glycan is from PDB entry 4tvp and the N156 glycan is from PDB entry 4dqo), contacts between the modeled glycans and the PGT128 antibody are expected (Fig. 2*b*). Together, these observations suggest that PGT128 accommodates high-mannose glycans at N137 and N156 by either making contacts that do not lead to any increase in affinity or by avoiding as much contact as possible. Nonetheless, if the glycans at these positions are hybrid or complex glycoforms, some interaction of the terminal sugars may be possible that would not be observed in our structure with oligomannose glycans. Ultimately, avoiding interaction with the highly heterogeneous complex N137 glycan should be beneficial for the breadth of PGT128, while interaction with the more conserved N156 glycan, which is thought to be mainly populated by a hybrid glycoform (Pancera *et al.*, 2013), should not adversely affect the breadth of PGT128. Recent advances in cryo-EM have made it the preferred method for analyzing glycans on HIV-1 Env (Lyumkis *et al.*, 2013) and on other heavily glycosylated proteins, where the proteins for structural studies can be expressed in 293T, 293F and other mammalian cells to produce oligomannose, hybrid and complex glycans on the expressed protein that would resemble those on the Env glycoprotein of the virus.

Finally, comparison with published SOSIP structures revealed a notable difference in the protein epitope of

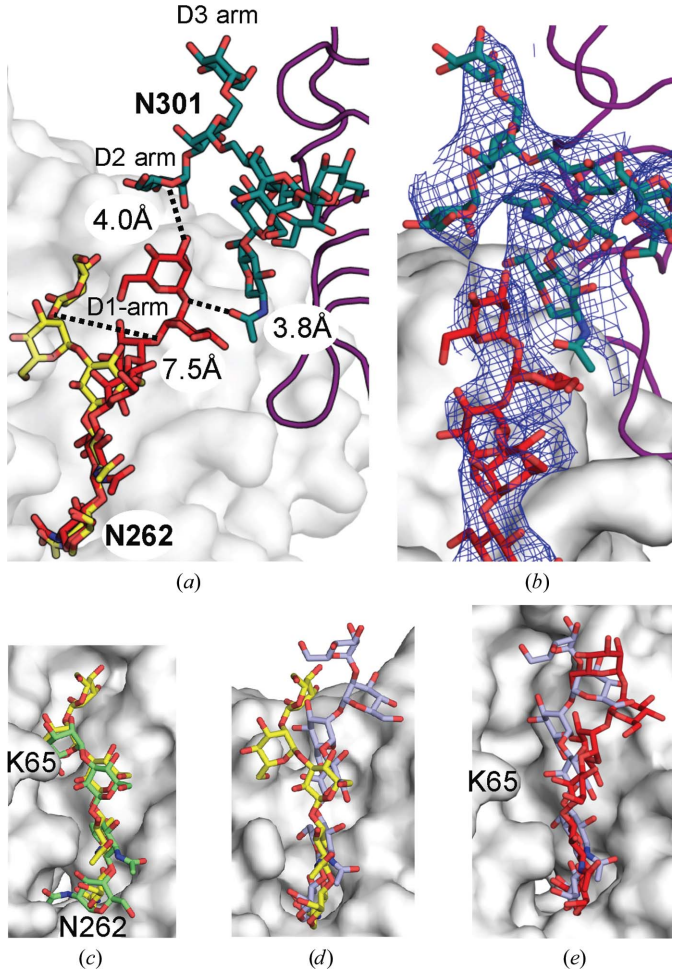


Figure 3
PGT128 indirectly interacts with the N262 glycan. (*a*) The interaction between the N301 glycan (ball-and-stick representation with green C atoms, red O atoms and blue N atoms) and the N262 glycan (ball-and-stick representation with red atoms) is shown. A purple ribbon diagram of PGT128 and a white surface representation of gp120 are shown to establish the context. For comparison, the N262 glycan from the PGT122-bound SOSIP structure (ball-and-stick representation with yellow C atoms, red O atoms and blue N atoms) is also displayed. Some of the interacting arms of the glycans (*e.g.* D1 and D2) are labeled. (*b*) The $2F_o - 2F_c$ electron-density map contoured at 1σ is shown for glycans N301 and N262. (*c*) Comparison of the N262 glycan from two previous structures of PGT122-bound SOSIP (PDB entry 4tvp, yellow; PDB entry 4nco, green). (*d*) Comparison of the N262 glycan from PGT122-bound SOSIP (PDB entry 4tvp, yellow) and a fully glycosylated CD4-bound gp120 core (PDB entry 4rqs, lilac). (*e*) Comparison of the N262 glycan from a fully glycosylated CD4-bound gp120 core (PDB entry 4rqs, lilac) and the PGT128-bound SOSIP (red). Residue Lys65 of gp120, which potentially interacts with the glycan, is labeled.

PGT128 and PGT122 (Pancera *et al.*, 2014). Both antibodies require contact with the 'GDIR' motif at the base of the V3 loop for binding, which is a highly conserved hotspot of interaction for the PGT121 and PGT128 family of antibodies (Garces *et al.*, 2014). PGT128 appears to induce a different conformation at this motif (Fig. 2*c*), highlighting the conformational adaptability of this region.

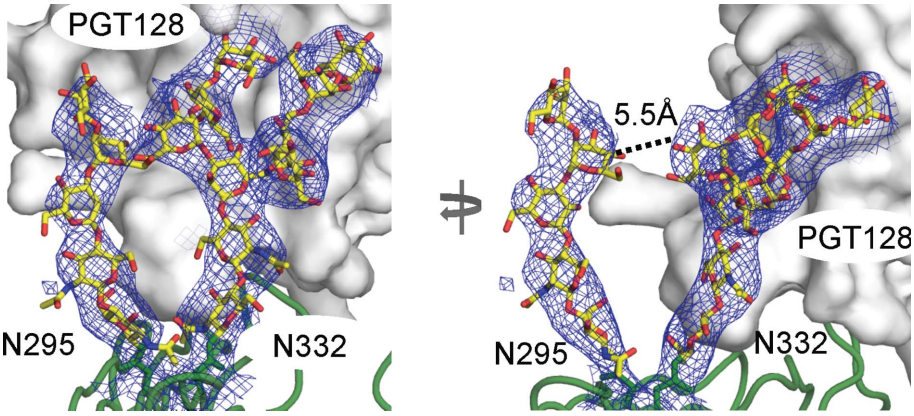
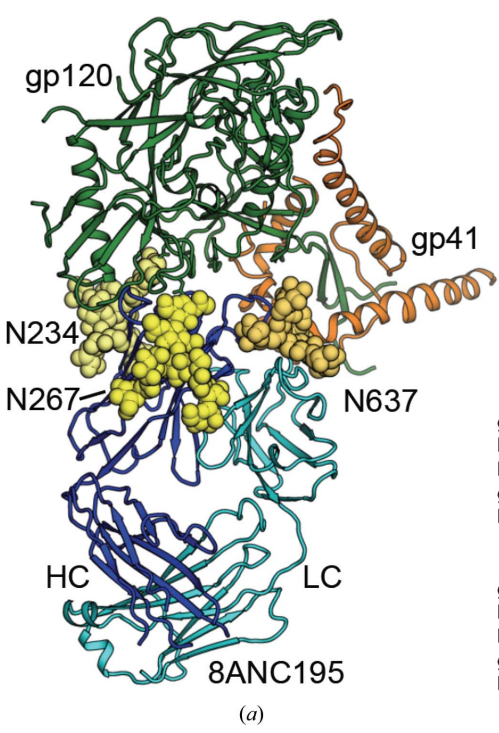


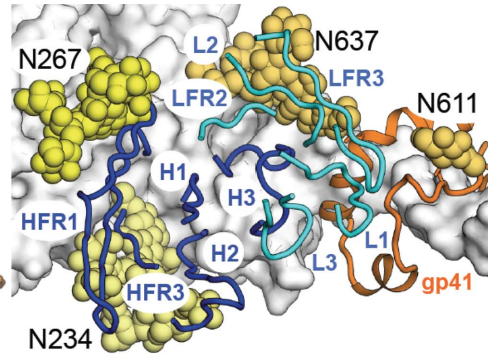
Figure 4
The N295 glycan of gp120 is poised to interact with PGT128 in the absence of the N332 glycan. The structures of the N332 and N295 glycans are displayed as ball-and-stick representations as in Fig. 1 within their $2F_o - 2F_c$ electron-density maps contoured at 1σ. The PGT128 protein is presented as a gray molecular surface interacting with the N332 glycan, while gp120 is displayed as a green ribbon diagram.

3.3. PGT128 makes allosteric interactions with the glycan shield

HIV-1 Env is densely glycosylated, suggesting that some changes in the glycan canopy would be expected when a bulky ligand such as an antibody covers an extensive surface on gp140. This phenomenon is observed for PGT128, which indirectly induces a conformation of the N262 glycan that was not previously observed. PGT128 binds to the N301 glycan, which in turn interacts with the N262 glycan (Fig. 3). Specifically, the D2 arm and the N-linked GlcNAc of the glycan at N301 contact the D1 arm of the N262 glycan (Fig. 3*a* and 3*b*). The observed PGT128-induced N262 glycan conformation is shifted 7–8 Å from its conformation in the PGT122-bound SOSIP structure. Since PGT122 does not engage the N301 glycan, its N262 glycan conformation may represent the native state. Indeed, both published PGT122-bound SOSIP structures display the same N262 glycan conformation (Julien *et al.*, 2013; Pancera *et al.*, 2014; Fig. 3*c*). By contrast, in the fully glycosylated core gp120 structure (Kong *et al.*, 2015) the N262 glycan displays a different orientation that is closer to that observed when PGT128 is bound (Figs. 3*d* and 3*e*). The core gp120 construct is heavily truncated at its N-terminus and therefore lacks a layer of protein that is present in the SOSIP trimer and that interacts with the N262 glycan. This difference between the constructs may explain the observed difference between the N262 glycan orientations. Notably, the N262 glycan is highly conserved and plays a critical role in gp120 folding (Kong *et al.*, 2015; Mathys *et al.*, 2014). Thus, this indirect glycan interaction might have global implications for gp120 function in addition to the observed structural changes to the glycan shield.



(*a*)



(*b*)

	HFR1	H1	HFR2	H2	HFR3	H3
gp120	74	111	0	101	254	274
N234	70	0	0	257	372	0
N276	254	0	0	0	37	0
gp41	0	0	0	0	0	197
N637	29	5	0	0	0	68

	LFR1	L1	LFR2	L2	LFR3	L3
gp120	0	0	0	0	0	0
N234	0	0	0	0	0	2
N276	0	0	12	0	0	0
gp41	0	223	0	60	15	9
N637	0	13	63	205	54	0

(*c*)

Figure 5
The 8ANC195 epitope on the SOSIP trimer. (*a*) A ribbon diagram of 8ANC195 bound to SOSIP gp140 is shown and colored as in Fig. 1. Interacting glycans are displayed as spheres. (*b*) Ribbon diagrams of interacting loops from 8ANC195 are shown on top of a white molecular surface of gp120 and an orange ribbon diagram of gp41. Glycans are shown as in (*a*). (*c*) Surface burial by loops on the 8ANC195 antibody of the gp120 and gp41 protein and glycans at N234, N276 and N637 are tabulated (Å^2). Abbreviations: HFR, heavy-chain framework region; LFR, light-chain framework region. Calculations were made in the *Molecular Surface Package* as described in §2.3.

The neighboring N295 glycan is also of considerable interest because PGT128 neutralization becomes dependent on it in

some HIV strains that lack the N332 glycan (Sok *et al.*, 2014). Here, we observe electron density for most of the D1 arm of the N295 glycan (Fig. 4). Although no direct interaction is observed between the N295 and N332 glycans, their D1 arms are separated by only 5 Å. Lack of the N332 glycan may therefore be compensated for by a simple shift of the N295 glycan to engage a similar binding site on PGT128. Further structural studies of PGT128 bound to SOSIP trimers lacking the N332 glycan would be required to fully understand this observation.

3.4. The 8ANC195 epitope

Although monomeric gp120 is able to engage the bNAb 8ANC195, it is clear from EM studies that a large portion of the 8ANC195 epitope is on gp41 (Scharf *et al.*, 2014). While the previously reported interactions with the gp120 protein and glycans are recapitulated in the SOSIP crystal structure, details of the interaction with the gp41 component of the epitope are now apparent (Fig. 5). Remarkably, although the antibody can bind to gp120 in the absence of gp41, 8ANC195 buries 943 Å^2 of the gp41 surface, constituting 34% of the surface area of the overall epitope. The 8ANC195 light chain plays a dominant role in gp41 recognition, burying 644 Å^2, while the heavy chain contacts gp41 mainly with its CDR3 (Figs. 5*a* and 5*b*). Only about half of the gp41 portion of the epitope is protein, encompassing a helix–turn–helix spanning residues 614–638 at the edge of heptad repeat 2 (HR2). The other half of the gp41 epitope is from glycan N637, which is 98% conserved in HIV-1 strains and is also recognized by PGT151 (Blattner *et al.*, 2014). The 8ANC195 light-chain CDR2, light-chain framework regions (LFR2 and LFR3) and heavy-chain CDR3 contact this glycan. The gp41 glycan N611, which is 99% conserved, is also near the epitope and may interact with the antibody, although electron density for this glycan was not visible in the crystal structure. Interactions with a complex glycan at this position might be observed using a more amenable methodology, such as cryo-EM.

To assess the contribution of gp41 protein and glycan components to antibody affinity, 8ANC195 binding was tested

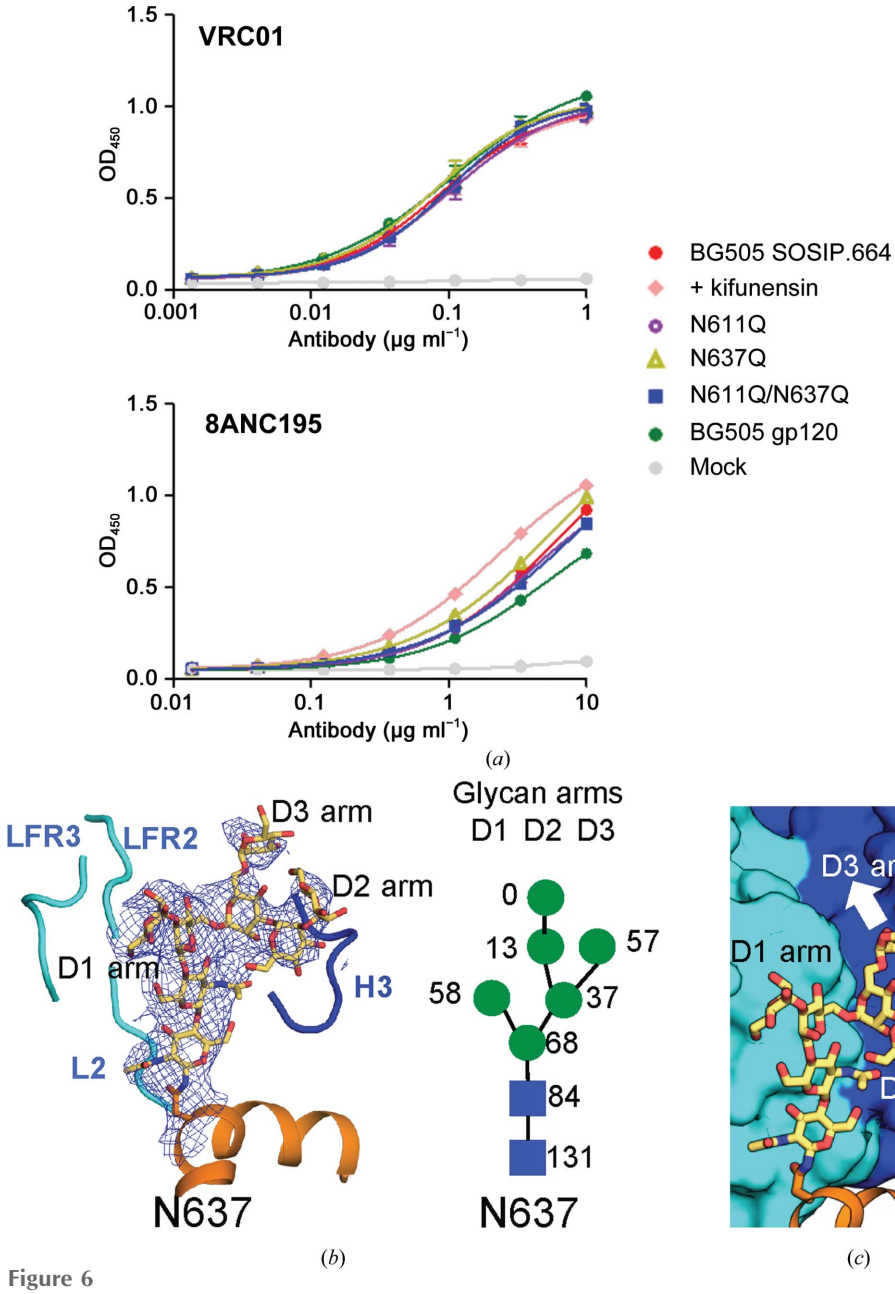


Figure 6
Interaction of 8ANC195 with gp41 glycans. (*a*) Representative ELISA curves of 8ANC195 binding to BG505 SOSIP.664 mutants expressed in HEK293T cell supernatant or BG505 SOSIP.664 expressed in the presence of kifunensine. The top panel displays ELISA curves for the CD4bs targeting VRC01 used as a control in the same experiment. (*b*) The structure of the N611 glycan is displayed in ball-and-stick representation and is colored as in Fig. 1 within its $2F_o - 2F_c$ electron-density map contoured at 1σ. Ribbon diagrams of contacting loops from 8ANC195, and the gp41 region to which the N637 glycan is attached, are displayed as in Fig. 5. To the right is a schematic of the glycan moieties in N637 with GlcNac moieties represented as blue squares and mannose moieties as green circles. The D1–D3 arms are labeled, and the per-moiety buried surface area (in Å^2) within the antibody–antigen interface is indicated. (*c*) The structure of the N637 glycan is displayed over a molecular-surface representation of the 8ANC195 antibody. The groove between the antibody heavy and light chains may accommodate variation in the D3 arm of the glycan and is highlighted by a white arrow.

against SOSIP gp140 without the N611 or N637 gp41 glycans, and against the gp120 monomer, all expressed in HEK293T cells (Fig. 6*a*). Binding was also tested against SOSIP gp140 produced in the presence of kifunensine, which results in glycoproteins decorated predominantly by $GlcNAc_2Man_9$ oligomannose glycans. The CD4 binding site (CD4bs) targeting bNAb VRC01, which should bind to all of these proteins equally, was used as a control (Fig. 6*a*). In line with previous studies (Scharf *et al.*, 2014), the presence of gp41 in the SOSIP gp140 trimer increased the 8ANC195 affinity slightly relative to that observed for monomeric gp120. Deletion of the N637 and N611 glycans did not significantly affect the 8ANC195 affinity, suggesting that the modest increase in binding affinity owing to interaction with gp41 is mainly from interactions with the protein. Finally, a subtle increase in affinity was observed following treatment with kifunensine, which may result in enhanced oligomannose glycan interactions observed in the structure. As expected, VRC01 binding remained the same for all of these envelope proteins.

Recent studies on the composition of glycoforms at specific glycan sites on HIV-1 Env indicate the gp41 glycans are highly heterogeneous (Pritchard *et al.*, 2015). Therefore, although the N611 and N637 glycans are nearly universally conserved, a wide range of glycoforms may be present at these positions. This finding may explain why 8ANC195 is only seen to interact with the core moieties of the N637 glycan, which are conserved across almost all glycoforms (Fig. 6*b*). Furthermore, the N637 glycan is positioned so that the D1 and D2 arms face away from 8ANC195, while the D3 arm extends towards a groove formed at the junction of the heavy and light chains of 8ANC195 (Fig. 6*c*). The antibody could therefore accommodate longer D1–D3 arms of the N637 glycan if the N637 site is occupied by more extended hybrid or complex glycans.

4. Conclusion

The crystal structure of the bNAbs PGT128 and 8ANC195 bound to the SOSIP gp140 trimer has enabled definition of the full epitopes for these important antibodies in the context of the native Env. The structure confirmed that PGT128 does not interact with the V1V2 glycans (N137 and N156) when they are oligomannose, while showing evidence for an indirect interaction with a neighboring glycan at position N262. The gp41 component of the 8ANC195 epitope is now elucidated and includes both protein residues and a glycan at position N637. Although these gp41 components of the antibody epitopes surprisingly do not contribute much to the full binding affinity, they represent an important surface that must be taken into account in generation of the germline antibody response, so that the correct antibody shape is selected for the accommodation of gp41 and associated glycans, and particularly for heterogeneity in N637 glycoforms. For PGT128, the selection of a particular germline antibody shape (in particular for the light chain) may also have been necessary to avoid heterogeneous contacts with the V1V2 glycans. An immunogen presenting only the hotspots of the antibody interaction, and lacking peripheral components of the epitope, may have superior properties in eliciting a response against the key conserved residues and glycans at the heart of the epitope. However, such an immunogen may elicit antibodies that clash with these peripheral elements, which can be (heterogeneous) glycans, variable loops and even gp41 in the case of 8ANC195. Presenting the immune system with the full antigenic complement of epitope components is therefore critical for vaccine design.

Acknowledgements

We are very grateful to M. Elsliger for computer support, H. Tien for crystallization screening and J. P. Verenini for manuscript formatting. X-ray data sets were collected on beamline 23-ID-B at the Advanced Photon Source (APS). Use of the APS was supported by the US Department of Energy, Basic Energy Sciences, Office of Science under contract No. DE-AC02-06CH11357. This work was supported by the HIV Vaccine Research and Design (HIVRAD) program (P01 AI832362 and P01 AI110657; IAW and RWS), the Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID UM1 AI100663; IAW), AI084817 (IAW), the International AIDS Vaccine Initiative Neutralizing Antibody Center (IAVI NAC) and the Collaboration for AIDS Vaccine Discovery (CAVD) funded by the Gates Foundation, and by the Joint Center of Structural Genomics (JCSG) funded by the NIH NIGMS, Protein Structure Initiative (U54 GM094586; IAW). RWS is the recipient of a Vidi grant from the Netherlands Organization for Scientific Research (NWO) and a Starting Investigator Grant from the European Research Council (ERC-StG-2011-280829-SHEV). A portion of this work was supported by an American Foundation for AIDS Research Mathilde Krim Fellowship in Basic Biomedical Research (LK). Author contributions: project design by LK and IAW, X-ray work and analysis by LK, MCD, FG, YH and RLS, and ELISA and mutagenesis studies by ATP, KS and RWS; the manuscript was written by LK, RLS, RWG and IAW. All authors were asked to comment on the manuscript. This is TSRI manuscript No. 29140. This work was partially funded by IAVI with the generous support from USAID, the Ministry of Foreign Affairs of the Netherlands and the Bill and Melinda Gates Foundation; a full list of IAVI donors is available at http://www.iavi.org. The contents of this manuscript are the responsibility of the authors and do not necessarily reflect the views of USAID or the US Government.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Blattner, C. *et al.* (2014). *Immunity*, **40**, 669–680.
Burton, D. R. *et al.* (2012). *Cell Host Microbe*, **12**, 396–407.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Connolly, M. L. (1993). *J. Mol. Graph.* **11**, 139–141.
Doores, K. J., Kong, L., Krumm, S. A., Le, K. M., Sok, D., Laserson, U., Garces, F., Poignard, P., Wilson, I. A. & Burton, D. R. (2015). *J. Virol.* **89**, 1105–1118.

Elsliger, M.-A., Deacon, A. M., Godzik, A., Lesley, S. A., Wooley, J., Wüthrich, K. & Wilson, I. A. (2010). *Acta Cryst.* F**66**, 1137–1142.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Garces, F., Sok, D., Kong, L., McBride, R., Kim, H. J., Saye-Francisco, K. F., Julien, J.-P., Hua, Y., Cupo, A., Moore, J. P., Paulson, J. C., Ward, A. B., Burton, D. R. & Wilson, I. A. (2014). *Cell*, **159**, 69–79.
Gelin, B. R. & Karplus, M. (1979). *Biochemistry*, **18**, 1256–1268.
Huang, J. *et al.* (2014). *Nature (London)*, **515**, 138–142.
Jardine, J. *et al.* (2013). *Science*, **340**, 711–716.
Julien, J.-P., Cupo, A., Sok, D., Stanfield, R. L., Lyumkis, D., Deller, M. C., Klasse, P.-J., Burton, D. R., Sanders, R. W., Moore, J. P., Ward, A. B. & Wilson, I. A. (2013). *Science*, **342**, 1477–1483.
Kong, L. *et al.* (2013). *Nature Struct. Mol. Biol.* **20**, 796–803.
Kong, L., Wilson, I. A. & Kwong, P. D. (2015). *Proteins*, **83**, 590–596.
Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C. & Detours, V. (2001). *Br. Med. Bull.* **58**, 19–42.
Kwong, P. D., Wyatt, R., Desjardins, E., Robinson, J., Culp, J. S., Hellmig, B. D., Sweet, R. W., Sodroski, J. & Hendrickson, W. A. (1999). *J. Biol. Chem.* **274**, 4115–4123.
Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J. & Hendrickson, W. A. (1998). *Nature (London)*, **393**, 648–659.
Lütteke, T. & von der Lieth, C. W. (2004). *BMC Bioinformatics*, **5**, 69.
Lyumkis, D., Julien, J.-P., de Val, N., Cupo, A., Potter, C. S., Klasse, P. J., Burton, D. R., Sanders, R. W., Moore, J. P., Carragher, B., Wilson, I. A. & Ward, A. B. (2013). *Science*, **342**, 1484–1490.
Martin, A. C. (1996). *Proteins*, **25**, 130–133.
Mathys, L., François, K. O., Quandte, M., Braakman, I. & Balzarini, J. (2014). *PLoS One*, **9**, e101181.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McLellan, J. S. *et al.* (2011). *Nature (London)*, **480**, 336–343.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pancera, M. *et al.* (2013). *Nature Struct. Mol. Biol.* **20**, 804–813.
Pancera, M. *et al.* (2014). *Nature (London)*, **514**, 455–461.
Pejchal, R. *et al.* (2011). *Science*, **334**, 1097–1103.
Pritchard, L. K., Vasiljevic, S., Ozorowski, G., Seabright, G. E., Cupo, A., Ringe, R., Kim, H. J., Sanders, R. W., Doores, K. J., Burton, D. R., Wilson, I. A., Ward, A. B., Moore, J. P. & Crispin, M. (2015). *Cell. Rep.* **11**, 1604–1613.
Sanders, R. W. *et al.* (2013). *PLoS Pathog.* **9**, e1003618.
Scharf, L., Scheid, J. F., Lee, J. H., West, A. P. Jr, Chen, C., Gao, H., Gnanapragasam, P. N. P., Mares, R., Seaman, M. S., Ward, A. B., Nussenzweig, M. C. & Bjorkman, P. J. (2014). *Cell. Rep.* **7**, 785–795.
Scheid, J. F. *et al.* (2011). *Science*, **333**, 1633–1637.
Sok, D., Doores, K. J., Briney, B., Le, K. M., Saye-Francisco, K. L., Ramos, A., Kulp, D. W., Julien, J.-P., Menis, S., Wickramasinghe, L., Seaman, M. S., Schief, W. R., Wilson, I. A., Poignard, P. & Burton, D. R. (2014). *Sci. Transl. Med.* **6**, 236ra63.
Stanfield, R. L., Ghiara, J. B., Ollmann Saphire, E., Profy, A. T. & Wilson, I. A. (2003). *Virology*, **315**, 159–173.
Vliegenthart, J. F. G., Dorland, L. & van Halbeek, H. (1983). *Adv. Carbohydr. Chem. Biochem.* **41**, 209–374.
Walker, L. M. *et al.* (2009). *Science*, **326**, 285–289.
Walker, L. M. *et al.* (2011). *Nature (London)*, **477**, 466–470.
Wei, X., Decker, J. M., Wang, S., Hui, H., Kappes, J. C., Wu, X., Salazar-Gonzalez, J. F., Salazar, M. G., Kilby, J. M., Saag, M. S., Komarova, N. L., Nowak, M. A., Hahn, B. H., Kwong, P. D. & Shaw, G. M. (2003). *Nature (London)*, **422**, 307–312.
Wyatt, R., Kwong, P. D., Desjardins, E., Sweet, R. W., Robinson, J., Hendrickson, W. A. & Sodroski, J. G. (1998). *Nature (London)*, **393**, 705–711.
Zhou, T. *et al.* (2010). *Science*, **329**, 811–817.