Crystal Structure of Human Antibody 2909 Reveals Conserved Features of Quaternary Structure-Specific Antibodies That Potently Neutralize HIV-1^{∇†§}

Anita Changela,¹ Xueling Wu,^{1‡} Yongping Yang,^{1‡} Baoshan Zhang,^{1‡} Jiang Zhu,^{1‡} Glenn A. Nardone,² Sijy O'Dell,¹ Marie Pancera,¹ Miroslaw K. Gorny,³ Sanjay Phogat,⁴ James E. Robinson,⁵ Leonidas Stamatatos,^{6,7} Susan Zolla-Pazner,^{3,8} John R. Mascola,¹ and Peter D. Kwong^{1*}

Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892¹; Research Technologies Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland 20852²; Department of Pathology, New York University School of Medicine, New York, New York 10016³; AIDS Vaccine Design and Development Laboratory, International AIDS Vaccine Initiative, Brooklyn, New York 11226⁴; Department of Pediatrics, Tulane University Medical Center, New Orleans, Louisiana 70012⁵; Seattle BioMed, Seattle, Washington 98109⁶; Department of Global Health, University of Washington, Seattle, Washington 98109⁷; and New York Veterans Affairs Medical Center, New York, New York 10010⁸

Received 8 November 2010/Accepted 8 December 2010

Monoclonal antibody 2909 belongs to a class of potently neutralizing antibodies that recognize quaternary epitopes on HIV-1. Some members of this class, such as 2909, are strain specific, while others, such as antibody PG16, are broadly neutralizing; all, however, recognize a region on the gp120 envelope glycoprotein that includes two loops (V2 and V3) and forms appropriately only in the oligomeric HIV-1 spike (gp120₃/gp41₃). Here we present the crystal structure of 2909 and report structure-function analysis with antibody chimeras composed of 2909 and other members of this antibody class. The 2909 structure was dominated by a heavy-chain third-complementarity-determining region (CDR H3) of 21 residues, which comprised 36% of the combining surface and formed a β -hairpin club extending ~ 20 Å beyond the rest of the antibody. Sequence analysis and mass spectrometry identified sites of tyrosine sulfation at the middle and top of CDR H3; substitutions with phenylalanine either ablated (middle substitution) or substantially diminished (top substitution) neutralization. Chimeric antibodies composed of heavy and light chains, exchanged between 2909 and other members of the class, indicated a substantial lack of complementation. Comparison of 2909 to PG16 (which is tyrosine sulfated and the only other member of the class for which a structure has previously been reported) showed that both utilize protruding, anionic CDR H3s for recognition. Thus, despite some diversity, members of this class share structural and functional similarities, with conserved features of the CDR H3 subdomain likely reflecting prevalent solutions by the human immune system for recognition of a quaternary site of HIV-1 vulnerability.

Identification of conserved regions accessible on the HIV-1 envelope and design of immunogens that elicit broadly neutralizing antibodies against these sites continue to be major challenges in the development of an effective HIV-1 vaccine. The HIV-1 viral spike—composed of three exterior gp120 subunits and three transmembrane gp41 subunits—is highly protected, but a limited number of these conserved regions exist on the spike, identified primarily by the broadly neutralizing antibodies that target them. One region is quaternary in nature and appropriately formed only on the assembled viral spike

(gp120₃/gp41₃). This region is targeted by a recently discovered (14) and fast expanding class of monoclonal antibodies (36, 40) that recognize epitopes with quaternary structural constraints, which are composed of portions of two gp120-variable loops, V2 and V3 (reviewed in reference 49). These quaternary structure-specific (or quaternary-specific) antibodies (also called quaternary-neutralizing epitope or “QNE” antibodies) are found in the sera of selected HIV-1-infected individuals who have broadly neutralizing serum antibodies (41); individual members of the class, however, vary greatly in their breadth of neutralization.

Initial evidence for the existence of quaternary-specific antibodies arose in simian/human immunodeficiency virus-infected rhesus macaques and HIV-1-infected chimpanzees (6, 9, 13). Characterization of polyclonal sera from these infected animals suggested the presence of antibodies targeting a conformational epitope involving the variable loop regions of the gp120 viral envelope.

Antibody 2909 was the first human monoclonal antibody against HIV-1 to be characterized as being specific for an

* Corresponding author. Mailing address: Vaccine Research Center, NIAID/NIH, 40 Convent Drive, Building 40, Room 4508, Bethesda, MD 20892-3027. Phone: (301) 594-8439. Fax: (301) 480-0274. E-mail: pdkwong@nih.gov.

‡ These authors contributed equally.

† Supplemental material for this article may be found at <http://jvi.asm.org/>.

[∇] Published ahead of print on 29 December 2010.

§ The authors have paid a fee to allow immediate free access to this article.

epitope dependent on the quaternary interaction of envelope glycoproteins (14). It was identified by direct screening for neutralization activity against a pseudovirus derived from strain SF162 of HIV-1. It recognizes a quaternary epitope on the surface of native virions and infected cells but does not bind soluble gp120/gp140 envelope proteins or cell surface-expressed gp120 monomers (14, 20). Competition analysis and virological assays indicate that the 2909 epitope includes portions of the V2 and V3 loops of gp120 (14, 16), with the V2-V3 elements originating either from within a gp120 monomer or between gp120 protomers in the trimer context. Mapping of 2909 recognition identifies a particular anomaly in its recognition (16); neutralization by 2909 depends on the presence of a rare lysine at position 160 in the V2 loop rather than the conserved N-linked site of glycosylation found at this position in most HIV-1 isolates (providing a residue-specific explanation for the neutralization specificity of 2909 for the SF162 virus, which contains this rare lysine).

Other strain-specific monoclonal antibodies like 2909 have been isolated from rhesus macaques infected with a chimeric simian/human immunodeficiency virus that contained an SF162 isolate-derived viral spike (SHIV_{SF162P4}) (36). These rhesus monoclonal antibodies exhibit properties similar to those of 2909 in their potent neutralization of SF162 and their recognition of V2-V3 only in the context of the functional viral spike (e.g., on virus particles) (36). Details from epitope mapping indicate that these rhesus antibodies and human antibody 2909 recognize overlapping epitopes, with some differences in requirements for V2 N-linked glycosylation (36).

The somatically related human monoclonal antibodies, PG9 and PG16, were also identified by a direct screen for neutralization (40). They target a quaternary-specific V2-V3 epitope, but unlike 2909, they neutralize an extraordinary 70 to 80% of circulating primary HIV-1 isolates and appear to have some reactivity for monomeric gp120 (40). Much of their increased breadth of neutralization arises from their ability to recognize an N-linked glycan at position 160 in the V2 loop, a motif which is found in greater than 90% of HIV-1 group M isolates (25).

Despite substantial differences in their neutralization breadth, antibodies 2909 and PG9/PG16 may be closely related. Notably, an N160K mutation in the V2 loop of typical primary HIV-1 isolates like YU2 and JR-FL can recover 2909 activity (16). Conversely, isolate SF162 can be converted to a PG9- and PG16-sensitive pseudovirus by the K160N mutation (40). Thus, a single N or K at position 160 appears to control much of the neutralization difference between 2909 and PG16. Together the results suggest that 2909 and PG9/PG16 antibodies recognize distinct immunotypes of a similar quaternary epitope.

To gain insight into how antibodies achieve recognition of this epitope, we determined the crystal structure of the antigen-binding fragment (Fab) of 2909 at a 3.3-Å resolution and compared this structure to the previously determined structure of PG16 (31, 33). Mutational analysis was used to confirm structural hot spots, and chimeric analysis of domain swaps between 2909 and other quaternary-specific antibodies was used to refine assessments of functional similarity. By identifying structural features—shared between 2909 and PG16 but otherwise highly uncommon in antibodies—the results provide

insight into conserved solutions by human antibodies for recognition of an important vaccine target on HIV-1.

MATERIALS AND METHODS

Expression and purification of 2909 antibody. The heavy and light chain sequences of the 2909 antibody (14) were codon optimized for mammalian expression, synthesized (Geneart), and individually subcloned into the pVRC8400 (CMV/R) mammalian expression vector (3). Signal peptide sequences used for the 2909 heavy and light chains were based on the corresponding signal peptides for the V-gene germ line precursor. To produce 2909 immunoglobulin, 2909 heavy and light chain plasmids were cotransfected into 1L of 293 FreeStyle cells with 293Fectin (Invitrogen). Cells were maintained in suspension culture at 8% CO₂, 37°C, 125 rpm. Five days after transfection, the supernatant was harvested, filtered through a 0.45-μm filter, and passed over a protein A affinity column. Bound antibody was eluted using immunoglobulin elution buffer (Pierce), immediately neutralized with 1 M Tris-Cl (pH 8.0), and dialyzed against phosphate-buffered saline (PBS).

To prepare the 2909 Fab, immunoglobulin was reduced and alkylated as described previously (22) and digested by Lys-C (Roche) at a 1:2,000 ratio (μg Lys-C per μg immunoglobulin) in 25 mM Tris-Cl (pH 8.5) and 1 mM EDTA for 4 to 6 h at 37°C. The digestion was quenched by addition of 1 mM *N*-α-p-tosyl-L-lysine chloromethyl ketone (TLCK), followed by purification over a Mono-Q (GE Healthcare) anion exchange column using a 0.1 to 1.0 M NaCl gradient in 25 mM Tris-Cl (pH 7.5) to separate Fab from other fragments. 2909 Fab was further purified by size exclusion chromatography (Superdex 200; GE Healthcare) and stored at 4°C.

Crystallization and data collection. 2909 Fab was concentrated to 10 to 15 mg/ml in 2.5 mM Tris-Cl (pH 7.5), 175 mM NaCl, and 0.01% sodium azide for crystallization. Initial crystals were thin needle clusters or plates that appeared at 20°C from robotic screening (Cartesian Honeybee robot; Digilab) in sitting drops (0.1 μl protein plus 0.1 μl precipitant) equilibrated against 20% polyethylene glycol (PEG) 8000 and 0.1 M Tris-Cl (pH 8.5). Crystallization conditions were further optimized using additive screening (Hampton Research). Diffraction-quality crystals were grown in hanging drops (0.5 μl Fab plus 0.5 μl precipitant) equilibrated against 17% PEG 8000, 0.1 M Tris-Cl (pH 8.5), 0.1 M NaCl, and 3% galactose. Prior to data collection, crystals were transferred in a single step (30 to 60 s) to a cryoprotectant containing 22% PEG 8000, 0.1 M Tris-Cl (pH 8.5), 0.1 M NaCl, 3% galactose, and 25% glycerol before being flash frozen in liquid nitrogen.

Diffraction data were collected at 100 K at beamlines BM-22 and ID-22 (SER-CAT) at the Advanced Photon Source, Argonne National Laboratory, and indexed and processed with HKL2000 (30).

Structure determination and refinement. The 2909 Fab structure was solved by molecular replacement with the software program PHASER (26) using diffraction data scaled to 4 Å. The constant and CDR loop-truncated variable domains of the PG16 Fab structure (Protein Data Bank [PDB] identifier [ID] 3LRS) (31) were used as individual search models. Successive solutions for individual domains led to initial placement of 3 complete Fab molecules in the asymmetric unit. The first three solutions were fixed, and a subsequent round of molecular replacement using the intact Fab molecule as a search model led to the placement of 3 additional Fab molecules. Model building was done using the software programs COOT (12) and O (18). Refinement was performed with the PHENIX program (1) using all diffraction data to 3.3 Å and a strategy that accounted for translation-libration-screw (TLS) vibrational motion and included group *B*-factor refinement with noncrystallographic symmetry (NCS) and strict geometric restraints. CDR loops and terminal residues for each chain were excluded from NCS restraints. Validation of the structure was carried out using the software program MolProbity (8) and the PDB validation server (<http://deposit.pdb.org/validate/>).

Structural homology of the 2909 CDR H3 loop. To search for structural homologs of the CDR H3 region, we submitted the region to the Dali server (15). No homologs of significant structural similarity were identified. We next constructed a database of proteins using the PISCES server (<http://dunbrack.fccc.edu/PISCES.php>) (42) with the following criteria: sequence identity of ≤90%, resolution of ≤3.0 Å, and *R* value of ≤1.0. The resulting database contained 16,938 proteins. We then structurally aligned the CDR H3 region on each of the proteins in the database using the TAlign software program (45) and processed the output using the following criteria: over 75% of the CDR H3 region (21 amino acids) is aligned onto the protein, the number of gaps in the alignment is ≤3, the length of alignment is ≤125% of the CDR H3 sequence, and the root mean square deviation (RMSD) of aligned residues is ≤4.0 Å. Note that in the current study, the alignment criteria have been applied to both the CDR H3

region and the proteins to be searched. By applying such criteria, any matches found are expected to have a length similar to that of 2909 with limited variations.

To investigate whether known antibodies contain similar CDR H3 structures, we constructed a second database that collects antibodies with long CDR H3 regions (≥ 16 amino acids), based on the SACS server (<http://www.bioinf.org.uk/abs/sacs/>) (2). The resulting database contained 99 structures. The 2909 CDR H3 loop was structurally aligned to CDR H3 of each antibody, and 23 of the closest homologs were selected for further comparison with respect to 2909. The detailed information of these antibodies and their structural alignment results are listed in Table S1 in the supplemental material.

Homology modeling of 2909 chimeras. For threading calculations, sequence alignments between 2909 and other quaternary-specific antibodies were generated using the ClustalW2 software program (24). Structural models of the 2909 chimeras were constructed using NEST (34), a model-building program based on rigid-body optimization. The threading scores were calculated using a normalized distance-scaled, finite ideal-gas reference (DFIRE) statistical potential (46).

Production of 2909 variants. PG16 heavy and light chain plasmids were generated as described previously (31). The heavy and light chain sequences for the rhesus monoclonal antibodies 2.3E, 2.2G, 2.5B, 1.8E, 1.6F, and LW10E (GenBank accession numbers HQ677009 through HQ677020) were either codon optimized for mammalian expression, synthesized (Geneart) and subcloned in frame with a murine immunoglobulin signal peptide sequence into the pVRC8400 expression vector, or placed in a pBR322-based expression vector for 1.6F (38). To produce chimeras between 2909 and other quaternary-specific antibodies, heavy and light chain plasmids were cotransfected into 293 Freestyle cells, expressed, and purified as described above for 2909 immunoglobulin.

CDR H3 swap variants were designed in which the entire CDR H3 sequence was exchanged between the 2909 and PG16 heavy chains. To generate the CDR H3 swap variants, the 2909 variable heavy chain sequence with PG16 CDR H3 (2909H_PG16CDRH3) and the PG16 variable heavy chain sequence with 2909 CDR H3 (PG16H_2909CDRH3) were synthesized (Geneart) and subcloned into the pVRC8400 expression vector. To further stabilize the PG16H_2909CDRH3 variant, site-directed mutagenesis (GM Biosciences and ACGT) was used to generate two variants of the PG16 light chain (PG16L-R96V and PG16L-D50V/L91W/R96V, which contained mutations to optimize interactions between the heavy and light chains).

Site-directed mutagenesis (ACGT) was used to create 2909 heavy chain variants in which potential sulfotyrosine sites in the CDR H3 were mutated to phenylalanine (2909H-Y^{100a}F and 2909H-Y^{100c}F).

All 2909 CDR H3 mutants and swap variants were expressed and purified in the same manner as described above for wild-type 2909 immunoglobulin. Purified immunoglobulin was exchanged into PBS, filtered through a 0.22- μ m filter, and concentrated to at least 1 mg/ml for neutralization assays.

ESI-MS measurements. For electrospray ionization mass spectrometry (ESI-MS) measurements, 2909 Fab was generated by sequential reduction, alkylation, and Lys-C digestion and purified by anion-exchange and size exclusion chromatography, as described above. ESI-MS data were collected using a Sciex 4000 Qtrap instrument at the Mass Spectrometry Lab, NIAID Research Technologies Branch.

Cells and HIV-1 *rev/*env plasmids. TZM-bl cells were obtained from the NIH AIDS Research and Reference Reagent Program, as contributed by John Kappes and Xiaoyun Wu. The full-length HIV-1 *rev/*env expression plasmids were obtained from the NIH AIDS Research and Reference Reagent Program.

Site-directed mutagenesis. Site-directed mutagenesis of HIV-1 *rev/*env plasmids was performed according to the instructions for the QuikChange II XL site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA). Primers containing the desired mutations were designed and synthesized. The wild-type HIV-1 *rev/*env plasmid was used as a template for amplification with each forward and reverse primer using PfuUltra high-fidelity DNA polymerase. The amplification product was digested with DpnI and then transformed into Top10 chemically competent cells (Invitrogen, Carlsbad, CA). Colonies were screened for the presence of the desired mutation by DNA sequencing. The entire HIV-1 *rev/*env region was sequenced for the final plasmid preparation (Qiagen Inc., Valencia, CA). To generate double mutations, plasmids containing a single mutation were used as a template.

Pseudovirus stocks and neutralization assays. HIV-1 Env pseudoviruses were prepared by transfecting 293T cells (6×10^6 cells in 50 ml growth medium in a T-175 culture flask) with 10 μ g of *rev/*env expression plasmid and 30 μ g of an *env*-deficient HIV-1 backbone vector (pSG3 Δ Envelope), using Eugene 6 transfection reagents (Invitrogen). Pseudovirus-containing culture supernatants were harvested 2 days after transfection, filtered (0.45 μ m), and stored at -80° C or in the vapor phase of liquid nitrogen. Neutralization was measured using HIV-1

Env pseudoviruses to infect TZM-bl cells as described previously (37, 44). Briefly, 40 μ l of pseudovirus was incubated for 30 min at 37° C with 10 μ l of serially diluted test antibody in duplicate wells of a 96-well flat-bottom culture plate. To keep assay conditions constant, sham medium was used in place of antibody in specified control wells. The pseudovirus input was set at a multiplicity of infection of 0.01 to 0.1, which generally results in 100,000 to 400,000 relative light units (RLU) in a luciferase assay (Bright Glo; Promega, Madison, WI). The antibody concentrations were defined at the point of incubation with pseudovirus supernatant. Neutralization curves were fit by nonlinear regression using a 5-parameter hill slope equation as previously described (37). The 50% inhibitory concentrations (IC₅₀) were reported as the antibody concentrations required to inhibit infection by 50%.

Structural alignments and figures. Structural superpositions were carried out with CCP4 software (10), and figures were prepared using the software program PyMOL (11).

Kabat numbering and residue nomenclature. We use the Kabat system of antibody residue numbering (19). To address a particular residue, we define residue type, heavy or light chain designation, and Kabat numbering. Thus, "Tyr^{H100a}" refers to tyrosine at position 100a in the heavy chain.

Protein structure accession number. Coordinates and structure factors for the 2909 crystal structure have been deposited in the Protein Data Bank (PDB) under accession code 3PIQ.

RESULTS

Production, crystallization, and structure solution of 2909

Fab. Transient transfection of plasmids encoding heavy and light chains of antibody 2909 generated ~ 80 mg of purified 2909 immunoglobulin per liter of cells. For the 2909 light chain produced by transient transfection, N-terminal sequencing showed cleavage to occur after the native N-terminal serine, likely a consequence of the signal peptide chosen. There was no difference, however, in neutralization activity by 2909 antibody produced from hybridomas or by transient transfection. Reduction, alkylation, and Lys-C proteolytic generation of the 2909 Fab, meanwhile, gave overall yields of 0.4 mg of Fab per mg of immunoglobulin. Crystals of 2909 Fab typically appeared a few days after setup and grew as hexagonal plates to maximum dimensions of 0.20 by 0.20 by 0.03 mm within 2 weeks.

Crystals were radiation sensitive, and diffraction intensities dropped off substantially after 5 $\text{\AA}$. Nonetheless, a complete trigonal data set to 3.3 $\text{\AA}$ was collected with 0.5° oscillations by using a 0.05-mm pinhole and translating the crystal to a fresh unexposed position every $\sim 25^\circ$ of data collection.

Molecular replacement with the previously determined PG16 structure (PDB ID 3LRS) (31) showed the space group to be P3₂21, with unit cell dimensions of $a = b = 180.4$ $\text{\AA}$ and $c = 222.5$ $\text{\AA}$, and to contain 6 Fab molecules in the asymmetric unit. Initial rigid body and TLS refinement in PHENIX (1) yielded an R factor of 40% ($R_{\text{free}} = 42\%$) and produced electron density maps that clearly showed the missing CDR loop regions. Final refinement to 3.3 $\text{\AA}$ with PHENIX (1) yielded an R factor of 24% ($R_{\text{free}} = 30\%$). Data collection and refinement statistics are shown in Table 1.

Overall structure of 2909 Fab. The six copies of 2909 Fab in the asymmetric unit refined to nearly identical conformations. Pairwise superposition of the Fab molecules yielded root-mean-square-deviation (RMSD) values ranging between 0.5 and 1.1 $\text{\AA}$ for all atoms (see Fig. S1 in the supplemental material). Several residues at the CDR H3 loop apex (residues 100 to 100h) were disordered in two of six copies of the Fab, and residues 127 to 133 in the heavy chain constant region were disordered in all of the molecules in the asymmetric unit; these residues were excluded from the final model. The struc-

TABLE 1. Data collection and refinement statistics

Parameter	Value(s)
Data collection	
Resolution (Å) ^a	50–3.3 (3.36–3.3)
Wavelength (Å)	1.0
Space group	P3 ₂ 21
Molecules per ASU	6
Unit cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	180.4, 180.4, 222.5
α , β , γ (°)	90, 90, 120
Measured reflections	257,052
Unique reflections	49,978
Completeness (%)	81.5 (43.6)
<i>R</i> _{sym} (%) ^b	15.6 (44.9)
Redundancy	5.1 (2.8)
<i>I</i> / σ (<i>I</i>)	8.8 (2.0)
Refinement	
Resolution (Å)	22–3.3
No. of reflections	47,598
<i>R</i> factor (%) ^c	23.7
<i>R</i> _{free} (%) ^d	29.9
RMSD from ideal	
Bond lengths (Å)	0.003
Bond angles (°)	0.577
Avg <i>B</i> factor (Å ²)	115.6
Ramachandran favored (%) ^e	90.8
Ramachandran outliers (%) ^e	0.7

^a Numbers in parentheses represent values in the highest-resolution shell.
^b $R_{\text{sym}} = \sum |I| - \langle I \rangle / \sum I$, where *I* = observed intensity, and $\langle I \rangle$ = average intensity obtained from multiple measurements.
^c R factor = $\sum ||F_o| - |F_c|| / \sum |F_o|$, where $|F_o|$ = observed structure factor amplitude and $|F_c|$ = calculated structure factor amplitude.
^d *R*_{free}, *R* factor based on 5% of the data excluded from refinement.
^e Calculated with MolProbity (8).

tural analysis presented here is based on Mol1 (chains H and L for heavy and light, respectively), which exhibited the most clearly defined electron density and lowest overall *B* factors. The final Mol1 model contained heavy chain residues 1 to 126

and 134 to 214 and light chain residues 2 to 208; notably, the entire CDR H3 of Mol1 was reasonably defined at the main-chain level.

2909 CDR H3 loop. The structure of the 2909 Fab revealed an antigen-combining site dominated by a protruding CDR H3 subdomain that extends ~20 Å beyond the rest of the combining loops and comprises almost 40% of the combining region surface area (Fig. 1). As defined by Kabat nomenclature and numbering (see Materials and Methods) (19), CDR H3 of 2909 is 21 residues in length, roughly 50% longer than the medium length of human CDR H3s. When measured from tip to base, the CDR H3 loop is ~30 Å in length (Fig. 2). Contiguous electron density was observed for much of the loop backbone, although many side chains, particularly at the apex of the loop, are not well defined (Fig. 2A). The CDR H3 loop is stabilized at its base with interactions from other heavy and light chain residues, while ~70% of the loop is solvent accessible (Fig. 2B).

Within the crystallographic asymmetric unit, the tip of the CDR H3 loop for any given Fab is found in close proximity to one or two other neighboring Fab molecules. Despite slightly different crystal packing environments, the structure of the CDR H3 loop is nearly identical in each of the four Fab molecules that exhibit an ordered loop, with pairwise RMSD values ranging from 0.3 to 0.8 Å for Cα atoms. The overall structure of the loop did not appear to be significantly influenced by any direct contacts between loop residues and neighboring molecules within the crystalline environment.

The base of the CDR H3 loop forms a bulged conformation stabilized by hydrogen bonding and stacking interactions with other CDR loop residues (Fig. 2B). Asp^{H95} is at the center of the bulge and interacts with His^{H35} from the CDR H1 region. Tyr^{L50} extends from the CDR L2 to hydrogen bond with Asp^{H98}. Arg^{H56} from the CDR H2 loop represents one of the few somatic mutations present in the 2909 gene (see Fig. S2 in

FIG. 1. Overall structure of 2909 Fab. The 2909 crystal structure reveals a combining region dominated by a protruding, acidic CDR H3 loop. (A) Ribbon representation of the 2909 Fab structure is shown, with heavy and light chains colored in blue and green, respectively. The CDR H3 loop is highlighted in red, while other CDR loops (as defined by Kabat [19]) are depicted in yellow. (B) Surface representation of the 2909 Fab is shown in the same orientation as in panel A (left) or rotated 180° about the *y* axis (right). The surfaces are colored by electrostatic potential (−10 to +10 kT/e), with positively and negatively charged regions shown in blue and red, respectively.

FIG. 2. Details of 2909 CDR H3. 2909 CDR H3 is a 21-residue loop forming a β -hairpin structure. The loop is stabilized at the base in a bulged conformation by other CDR loop residues. Most of the CDR H3 loop extends out from the combining site and is solvent accessible. (A) $2F_o - F_c$ electron density contoured at 1σ is shown around 2909 CDR H3. Heavy and light chains are colored blue and green, respectively, while the CDR H3 loop is shown in pink stick representation. The conformations of side chains at the tip of the CDR H3 loop are not clear due to weak density in this region. (B) Hydrogen bonds stabilizing the 2909 CDR H3 loop are drawn as dotted lines. Coloring is the same as in panel A, with additional stabilizing residues from heavy and light chains shown in gray stick.

the supplemental material) and stabilizes the CDR H3 loop through π -cation interactions with the phenyl ring of Tyr^{H100i} and hydrogen bonding with Tyr^{H100k}. The central portion of the loop is maintained in an antiparallel β -sheet structure composed of the residues Asp^{H98}, Ser^{H99}, and Asp^{H100} from the ascending strand which hydrogen bonds to Ser^{H100h}, Tyr^{H100i}, and Phe^{H100j} on the opposite strand (Fig. 2B). A five-residue turn forms the tip of the CDR H3 loop. The electron density was weak for residues at the loop apex, and high B factors suggested inherent flexibility for this region.

Structural homology of 2909 CDR H3. In a search of the Protein Data Bank, 2,272 entries were found to contain structural homologs of 2909 CDR H3, suggesting that β -hairpins of this geometry and length are relatively common in protein structures. A notable feature of 2909 CDR H3 is that much of the loop is exposed (67.1%). Histogram analysis of the solvent accessibility of all structural homologs showed a broad distribution, with two-thirds of the structures falling into the lower half of the distribution (Fig. 3A). Only 22 loop structures have $\geq 60\%$ of their surface exposed. Among these 22 structures, 20 formed complexes by using the loop region to pair with or to pack against β -sheets of other subunits in the complex. The two remaining monomeric structures, 1S7I and 1XJC, have unknown functions. We also examined the sequence similarity of structural homologs to 2909 CDR H3. No structural homolog had a sequence identity to 2909 CDR H3 of greater than 30%, with two-thirds of all homologs sharing less than 15% identity (Fig. 3B). Overall, the exposed nature of the long β -hairpin makes 2909 CDR H3 an unusual structural motif in the general population of protein structures.

Additionally, CDR H3 of 2909 was compared to long CDR H3 regions found in other antibodies, using a database based on the SACS server (<http://www.bioinf.org.uk/abs/sacs/>) (2). The 2909 CDR H3 loop was structurally aligned to the CDR

H3 of each antibody in the database (the closest homologs at least 16 residues in length are given in Table S1 in the supplemental material), with 12 out of the 23 antibodies listed being anti-HIV-1 antibodies. One of the closest structural homologs identified was CDR H3 of anti-V3 antibody 447-52D, which gave an average RMSD of 1.76 Å over C α atoms and an alignment of 19 (out of 20) residues (Fig. 3C). The second-best match was the anti-gp41 antibody Z13e1, which showed an RMSD of 1.63 Å over C α atoms for 17 aligned residues. Despite the structural similarity, 447-52D and Z13e1 use their CDR H3 loops to interact with their respective epitopes in different ways: 447-52D uses one of its CDR H3 strands to pair with the V3 loop primarily through backbone hydrogen bonds and a few side chain interactions, while the Z13e1 CDR H3 packs against the gp41 loop through side chain interactions. In light of the high degree of structural homology between 2909 and 447-52D and the fact that the 2909 epitope includes the V3 loop, it is possible that part of the 2909 CDR H3 loop interacts with V3 in a way similar to that of 447-52D (although the alternative mode of recognition used by Z13e1 indicates that other possibilities cannot be excluded).

Anionic character of CDR H3 loop and tyrosine sulfation.

The 2909 CDR H3 loop sequence is tyrosine rich within an overall acidic context (Fig. 1B). This combination of tyrosine residues and anionic character is similar to the sequence signature recognized by human sulfotransferases, which introduce O sulfation of tyrosine as a posttranslational modification (28). The precise sequon for O sulfation by sulfotransferase is not known, but reasonable predictions have been established with neural-network-based software programs, such as Sulfinator (27) and Sulfosite (5). With 2909, both programs predict sites of tyrosine sulfation, at Tyr^{H100a} and Tyr^{H100c}, at the middle and tip of CDR H3, respectively. Although additional sites are predicted by Sulfinator, both Tyr^{H100a} and Tyr^{H100c}

FIG. 3. Structural homology of 2909 CDR H3 region. Sixty-seven percent of the long β -hairpin structure of the 2909 CDR H3 loop is exposed to solvent, making it a rare structural motif among all protein structures. Structural homologs to the CDR H3 show <30% sequence identity. Several anti-HIV-1 antibodies have structurally similar CDR H3 regions, including 447-52D and Z13e1. (A) Two thousand two hundred seventy-two structural homologs were identified from a database of 16,938 proteins and analyzed in a statistical manner. The distribution of solvent accessibility of structural homologs is shown, with the value of 2909 CDR H3 marked by a red arrow. Homologs range in length from 15 to 25 residues. (B) The distribution of sequence identity of structural homologs is shown with respect to the 2909 CDR H3. (C) A structural comparison of the 2909 CDR H3 loop with two structural homologs, 447-52D and Z13e1, is shown by superposition of the heavy chain variable domains using LSQKAB (10). 2909 heavy and light chains are colored in blue and green, respectively, while 447-52D and Z13e1 are gray. The CDR H3 regions of 2909, 447-52D, and Z13e1 are highlighted in red, orange, and yellow, respectively.

are solvent exposed, increasing the likelihood that they are O sulfated (see Table S2 in the supplemental material). Electron density was not well defined for these residues, especially for Tyr^{H100a}, where the entire phenol ring appears to be conformationally disordered, thereby preventing crystallographic confirmation of O sulfation (Fig. 2A). Nonetheless, the presence of two sites of sulfotyrosine modification on the 2909 heavy chain was confirmed by electrospray ionization mass spectrometry (ESI-MS) analysis (Table 2 and Fig. 4).

We decided to test the functional role of Tyr^{H100a} and Tyr^{H100c} by mutating the tyrosine to phenylalanine at both sites and assaying the mutants for neutralization (Table 3). Mutational substitution of Phe for Tyr^{H100a} in the middle of the CDR H3 ablated neutralization activity in all 5 isolates tested that were sensitive to 2909. Substitution of Phe for Tyr^{H100c} at the tip of the loop ablated neutralization activity against SF162 and YU2.N160K and diminished neutralization (20- to 25-fold

TABLE 2. ESI-MS analysis of 2909 Fab

Protein	Theoretical mass (Da) ^a	Exptl mass	Difference in mass ^b	No. of sulfotyrosines
2909 Fab, heavy chain	25,343	25,507	+164	2
2909 Fab, light chain	22,848	22,851	+3	0
Lysozyme ^c	14,306	14,307	+1	0

^a Theoretical mass for 2909 is calculated based on expected amino acid sequence of the Fab generated after Lys-C digest, with additional mass (+57 Da) added to both heavy and light chain masses due to reduction and alkylation by iodoacetamide of the single interchain disulfide. Calculated mass of light chain is based on loss of native N-terminal serine, a consequence of alteration of the signal peptide sequence in the immunoglobulin expressed by transient transfection from synthetic gene (see Methods).

^b Difference between theoretical and experimental masses. Presence of a sulfogroup corresponds to a +80 change in mass.

^c Theoretical and experimental masses for the hen egg white lysozyme standard are based on the native form, which includes 4 disulfide bonds.

FIG. 4. Electrospray ionization mass spectrometry (ESI-MS) data for 2909 Fab. ESI-MS intact mass measurements of 2909 Fab are consistent with two sites of tyrosine sulfation in the heavy chain. Deconvoluted ESI mass spectra are shown for 2909 Fab heavy and light chains (A) or a hen egg white lysozyme standard (B).

increase in IC₅₀) against three other 2909-sensitive pseudoviruses containing the N160K mutation (Table 3). The results indicate a critical role for these tyrosines, both of which are located on the heavy chain-facing side of the CDR H3.

Comparison of antibodies 2909 and PG16. Unlike antibody 2909, antibody PG16 is broadly reactive and can neutralize 70 to 80% of HIV-1 isolates (40). Nevertheless, 2909 and PG16 appear functionally related: mapping reveals recognition of a similar V2-V3 epitope, distinguished by a requirement for Lys160 in the V2 loop for 2909 and Asn160 for PG16. Antibodies 2909 and PG16 use different light chains (IGLV3-21*01 and IGLV2-14*01, respectively) but share 68% sequence identity in their variable heavy chain genes, both of which are

derived from VH3 gene precursors (IGHV3-43*01 and IGHV3-33*05 for 2909 and PG16, respectively) and the IGHJ6*03 gene (Fig. 5A).

To determine whether the sequence and functional similarities between 2909 and PG16 correlate with the use of common structural elements, we compared the 2909 structure to the recently determined crystal structure of PG16 Fab (31, 33). Structural comparison of the two antibodies revealed common features. Both 2909 and PG16 have long, acidic CDR H3 loops which form subdomains that protrude from the antibody surface (Fig. 5B) and comprise significant portions of the combining surface: in 2909, CDR H3 contributes 36% of the CDR surface area; in PG16, CDR H3 contributes 42%. Moreover, CDR H3s in both PG16 (33) and 2909 are tyrosine sulfated.

PG16 accommodates an additional 7 amino acids at the loop tip, giving it an overall hammerhead or axe shape (31, 33), rather than the headless axe or club shape observed in the 2909 structure (Fig. 5B). In both cases structural homologs to the CDR H3 were identified, and in both cases homology analysis indicated that a high degree of solvent exposure was unusual.

Critical residues in the CDR H3 domain for 2909 and PG16 are found at approximately the same distance from the rest of the combining site surface in both antibodies (Fig. 5B). For PG16, the most critical residue identified by mutational analysis is Asp^{H100i}; replacement of this residue by alanine greatly diminishes PG16 neutralization activity (33). For 2909, the potentially tyrosine sulfated Tyr^{H100a} is critical for activity. Thus, critical negative charges reside in highly similar places in the 2909 and PG16 paratopes (Fig. 5B).

A closer look at the electrostatics surrounding the antigen-combining site revealed an acidic patch, including CDR H3 for

TABLE 3. IC₅₀ neutralization values for 2909 CDR H3 tyrosine mutants

Clade	Virus ID	IC ₅₀ (μg/ml) of ^a :		
		2909	2909-Y ^{H100a} F	2909-Y ^{H100c} F
A	Q23.17	>50	>50	>50
	Q23.17.N160K	0.048	>50	1.0
B	SF162	0.084	>50	>50
	SF162.K160N	>50	>50	>50
	YU2	>50	>50	>50
	YU2.N160K	0.312	>50	>50
	JR-FL.E168K	>50	>50	>50
	JR-FL.N160K/E168K	0.001	>50	0.023
C	ZM233.6	>50	>50	>50
	ZM233.6.N160K	0.153	>50	3.22

^a Boldface indicates IC₅₀ of <50 μg/ml; italics indicate IC₅₀ of >50 μg/ml.

FIG. 5. Sequence and structural comparison of 2909 with PG16. 2909 and PG16 are both from VH3 gene families. 2909 and PG16 both have critical residues in similar positions in their CDR H3 sequences, and these residues are located at approximately the same distance from the combining surface in both antibodies. Acidic CDR H3 regions for 2909 and PG16 are found on opposing faces of each antibody. (A) Sequence alignment of 2909 and PG16 heavy chains is shown, with identical residues shaded in gray. Tyrosine sulfation in PG16 CDR H3 and predicted sites of sulfation in 2909 are highlighted in magenta. The Kabat numbering scheme is used for CDR loop definition. (B) A comparison of 2909 and PG16 CDR H3 loop structures is shown. The variable domains of 2909 and PG16 were superimposed (over $\text{C}\alpha$ atoms, excluding the CDR H3 loop) and are shown on top. The bottom panel shows the same orientation for PG16 while 2909 is flipped, such that the 2909 V_H is aligned with PG16 V_L and 2909 V_L is aligned with PG16 V_H . The PG16 model is based on PDB ID 3MUG (33). PG16 is colored gray, with its CDR H3 loop depicted in tan. Sulfated $\text{Tyr}^{\text{H100h}}$ and $\text{Asp}^{\text{H100i}}$ are important residues for PG16 activity and are drawn in gray stick. 2909 is colored as in Fig. 1A, with $\text{Tyr}^{\text{H100a}}$ and $\text{Tyr}^{\text{H100c}}$ shown in red stick. (C and D) A close-up view of the combining sites from 2909 (C) or PG16 (D). The surfaces are colored by electrostatic potential (-10 to $+10$ kT/e), with positively and negatively charged regions shown in blue and red, respectively. CDR loops are labeled accordingly.

both PG16 and 2909 (Fig. 5C and D). Interestingly, the negatively charged surfaces were found on opposing faces of each antibody, with CDR L2 residues making additional anionic contributions in 2909 and CDR H2 residues contributing in PG16.

If 2909 and PG16 were superimposed based on electrostatic similarities, the heavy chain of 2909 would align with the light chain of PG16 and vice versa (Fig. 5B, bottom panel). Notably, such a superposition would also orient the functionally critical $\text{Asp}^{\text{H100i}}$ in PG16 and the functionally critical $\text{Tyr}^{\text{H100a}}$ on similar faces of their respective CDR H3s. For PG16, the tyrosine-sulfated residue ($\text{Tyr}^{\text{H100h}}$) is not so critical for PG16 function, with replacement by phenylalanine at this site in PG16 resulting in only a 2.2-fold increase in the IC_{50} (33). The location of negative charge at a particular place in the paratope may thus be more critical than the precise character of the side chain.

Chimeric analysis of antibodies 2909 and PG16. In the absence of structural information for quaternary-specific anti-

bodies complexed with their epitopes, we used data from antibody chimeras to gain insight into how the CDR H3, heavy chain, and light chain elements contribute to the 2909 paratope. In this manner, regions within 2909 could be probed individually within the context of a related quaternary-specific antibody. Differences in sequence could be mapped out on models of chimeric antibodies that complement functionally to delineate functional hotspots.

To assess the individual contributions of heavy and light chain elements to 2909 and PG16 activity, heavy and light chains were swapped and the chimeric antibodies were assayed for neutralization against 2909-sensitive and PG16-sensitive pseudoviruses. Both chimeric antibodies could be produced, as verified by SDS-PAGE analysis of heavy and light chains, although the 2909H/PG16L swap expressed at low levels, suggesting that structural differences between the 2909 and PG16 light chains may affect the dimer interface. Neither chimeric antibody was active against SF162 and N160K mutant pseudoviruses or any PG16-sensitive pseudovirus (Table 4).

TABLE 4. IC₅₀ neutralization values^a for 2909/PG16 chimeras

Clade	Virus ID	CDR H3		PG16	2909H/PG16L	PG16H/2909L	2909H_PG16CDRH3/2909L ^b	2909H_PG16CDRH3/PG16L	PG16H_2909CDRH3/2909L
		Light	Heavy						
A	Q23.17	>50	>50	<0.003	>50	>50	>50	>50	>50
	Q23.17.N160K	0.048	>50	>50	>50	>50	>50	0.592	>50
	SF162	0.084	>50	>50	>50	>50	>50	>50	31.1
	SF162.K160N	>50	>50	<0.001	>50	>50	>50	>50	>50
B	YU2	>50	>50	0.064	>50	>50	>50	>50	>50
	YU2.N160K	0.312	>50	>50	>50	>50	>50	>50	>50
	JR-FL.E168K	>50	>50	<0.003	>50	>50	>50	>50	>50
	JR-FL.N160K/E168K	0.001	>50	>50	>50	>50	>50	>50	1.14
C	ZM233.6	>50	>50	0.001	>50	>50	>50	>50	>50
	ZM233.6.N160K	0.153	>50	>50	>50	>50	>50	>50	>50
Non-HIV	SIVmac251.30	>50	>50	>50	>50	>50	>50	>50	>50
	MuLV	>50	>50	>50	>50	>50	>50	>50	>50

^a Expressed in µg/ml. Boldface indicates IC₅₀ of <50 µg/ml; italics indicate IC₅₀ of >50 µg/ml.

^b The PG16H_2909CDRH3/PG16L chimera did not express and could not be tested.

To test the role of CDR H3 in neutralization, we swapped the CDR H3 sequence between 2909 and PG16 and assayed the chimeras for activity. A chimera composed of 2909 with the PG16 CDR H3 loop (2909H_PG16CDRH3/2909L) was not active against any isolate tested (Table 4). A PG16 variant containing 2909 CDR H3 (PG16H_2909CDRH3/PG16L) did not express. Threading analysis and structural modeling of the chimera suggested that residues from the PG16 light chain might clash with 2909 CDR H3, thereby affecting the dimer interface and CDR H3 conformation. To further stabilize the PG16H_2909CDRH3 variant, site-directed mutagenesis was used to generate two variants of the PG16 light chain (PG16L-R96V and PG16L-D50V/L91W/R96V) that contained mutations to optimize interactions between the heavy and light chains. The redesigned chimeras also did not express. When the CDR H3 loop from each antibody was paired with its native light chain, however, neutralizing activity was recovered, although at a considerably reduced level. For example, when the PG16 heavy chain containing the 2909 CDR H3 loop was paired alongside the 2909 light chain (PG16H_2909CDRH3/2909L), the chimera exhibited weakened activity against select 2909-sensitive pseudoviruses (Table 4, last column). Similarly, when the 2909 heavy chain containing the PG16 CDR H3 was paired with the PG16 light chain (2909H_PG16CDRH3/PG16L), PG16-like activity was recovered for a single isolate, albeit at ~1000-fold reduced potency (Table 4, next to last column). Although the data suggest a role for the light chain in antigen recognition, they do not indicate whether the light chain is contributing primarily a functional role or a structural role in maintaining CDR H3 loop conformation.

Relationship of antibody 2909 to quaternary-specific rhesus monoclonal antibodies. 2909-like rhesus antibodies have been isolated that are strain specific in their activity toward SF162 and recognize V2-V3 quaternary neutralizing epitopes, which are present only on HIV-1 virions and infected cells (36). To further investigate the functional similarity to 2909, the rhesus antibodies were assayed for neutralization activity against two 2909-sensitive pseudoviruses. As expected, all rhesus antibodies were extremely potent against SF162 (Table 5). While 2909 can also neutralize YU2 containing an N160K substitution (16), none of the rhesus antibodies showed activity against the mutant YU2 pseudovirus (Table 5).

Next, chimeric antibodies were used to gain insight into the degree of similarity between 2909 and the quaternary-specific rhesus antibodies. Heavy and light chain swaps generated between 2909 and the rhesus antibodies were tested against 2909-sensitive isolates. All of these chimeras were expressed and could be purified, thereby suggesting that any sequence differences between 2909 and rhesus genes did not prevent chain association or have a major effect on the structural integrity of the antibody. Chimeras containing a rhesus heavy chain paired with the 2909 light chain retained some level of neutralizing activity against SF162 but were not active against the YU2.N160K pseudovirus (Table 5). For three of the six rhesus-H/2909-L chimeric antibodies, little loss of activity was noted, while the remaining three showed an IC₅₀ increase of 2,000- to 17,000-fold. Notably, the 2.5B heavy chain could be combined with the 2909 light chain with virtually no loss in neutralization potency against SF162. The immunoglobulin gene usage and

TABLE 5. IC₅₀ neutralization values for 2909/rhesus antibody chimeras

Antibody	IC ₅₀ (μg/ml) against:					
	SF162.LS			YU2.N160K		
	Wild-type	Rhesus-H/2909-L	2909-H/Rhesus-L	Wild-type	Rhesus-H/2909-L	2909-H/Rhesus-L
2909	0.084	NA ^b	NA	0.312	NA	NA
2.3E ^a	2.63	11.5	>50	>50	>50	>50
2.2G ^a	<0.003	6.12	>50	>50	>50	>50
2.5B	<0.003	0.003	>50	>50	>50	>50
1.8E	0.003	50	>50	>50	>50	>50
LW10E	<0.003	50	>50	>50	>50	>50
1.6F	<0.003	0.012	>50	>50	>50	>50

^a 2.3E and 2.2G are derived from the same clone.

^b NA, not applicable, where 2909 is a human antibody and does not apply to Rhesus-H or Rhesus-L.

sequence homology to 2909 are variable among all members of the rhesus group (see Appendix S1 in the supplemental material) and did not correlate with the degrees of activity exhibited by the different chimeras.

Chimeric antibodies comprising the 2909 heavy chain and a rhesus light chain were not active against any of the isolates tested. Several of the rhesus light chain sequences share ~90% sequence similarity to the 2909 light chain. To gain insight into the lack of function for chimeras composed of 2909 heavy chain and rhesus light chains, all of the chimeras were threaded onto the 2909 structure (see Table S3 in the supplemental material). Threading analysis suggested a potential explanation for the lack of complementation related to the heavy-light chain interface: residue 96 of the 2909 light chain is a valine, which can accommodate large side chains from the 2909 heavy chain. In contrast, rhesus antibodies have larger amino acids at this position in their light chain, which may be incompatible with the 2909 heavy chain.

DISCUSSION

The structure-function analyses presented here provide a framework for understanding how different molecular features contribute to various specificities of recognition by 2909 and other quaternary-specific antibodies. In addition, analysis of conserved molecular features—shared by members of the class but otherwise rare in antibodies and proteins in general—provides insight into structural requirements for antibody recognition of this quaternary-dependent site of HIV-1 vulnerability.

Conserved features of 2909, PG9, PG16, and rhesus quaternary-specific antibodies. The presence of a protruding CDR H3 loop is a striking feature of the 2909 antigen-combining surface. This feature is also observed in the PG16 crystal structure (Fig. 5B) and the modeled PG9 structure (31) and is consistent with threading of some of the rhesus antibodies (e.g., 2.5B) onto the 2909 structure. Although the structural motifs that represent CDR H3 (hammerhead or axe with PG9/PG16; club with 2909) are seen in other protein structures, the highly solvent-exposed nature of this motif as seen in these quaternary-specific antibodies appears to be generally rare.

The prediction of tyrosine sulfation is another shared feature of 2909, PG9, PG16, and several of the rhesus antibodies that is also uncommon in antibodies. Whether or not sulfation actually occurs has been demonstrated only with 2909, PG9,

and PG16 (32–33); independent of actual sulfation, the prediction of O sulfation highlights the high tyrosine content and acidic nature of CDR H3 of the quaternary-specific antibodies.

Finally, the functional dominance of CDR H3 is another shared feature of 2909, PG9, and PG16 that is somewhat uncommon in antibodies.

Differences between 2909 and other quaternary-specific antibodies. Details of the 2909 structure and functional data from antibody chimeras clarified the relationship of 2909 to other members of the quaternary-specific class of antibodies. Both 2909 and PG16 are highly dependent on the residue at position 160 in the V2 loop. A K160N mutation in SF162 and N160K mutations in selected pseudoviruses can allow the pseudovirus to switch between 2909 and PG16 sensitivities. Interestingly, studies with SHIV_{SF162}-infected rhesus macaques show that as the circulating virus evolves to escape autologous neutralizing antibodies, one of the first mutations found (appearing within the first 2 months of infection) is the K-to-N change at position 160 in V2 (21). Indeed, accumulating evidence indicates that the primary reason 2909 does not neutralize more broadly relates specifically to the N160K substitution, thereby suggesting that 2909 and PG16 recognize different immunotypes of the same epitope (Table 4). The structural comparison of 2909 and PG16, highlighted here, indicates that these antibodies use protruding CDR H3 structures with moderately different shapes and possibly differing heavy and light chain orientation in recognition of a similar quaternary-specific epitope.

Functional analysis of the 2909/rhesus chimeras suggests activity by the rhesus monoclonal antibodies is predominantly mediated by the heavy chain. Although the data from the PG16_2909CDRH3/2909L chimera implicate a role for the light chain in 2909 activity, this role could be mainly structural in maintaining the CDR H3 loop conformation. Thus, it is possible that 2909 activity is primarily heavy chain mediated, similar to that of the rhesus antibodies.

Insights into the quaternary-specific site of vulnerability. Do these antibody structures provide any insight into the recognized quaternary V2-V3 region on the trimeric envelope spike? Due to the protruding acidic nature of CDR H3 in 2909 and PG16, it is tempting to speculate that CDR H3 of these quaternary-specific antibodies may allow access to a recessed cationic site on the functional viral spike. Analysis of other anti-HIV-1 antibodies with various specificities and extended CDR H3 regions, such as antibodies 17b, 447-52D, 537-10D,

b12, F105 and X5, however, indicates that they do not use their loops to reach into a recessed site (4, 7, 17, 23, 39, 48). Rather, in all cases, these antibodies recognize conserved elements that are widely separated. For example, with 17b and X5, recognition occurs at a hydrophobic site on CDR H2 at the surface of the combining loops and with an extended CDR H3, which reaches toward V3 (17, 23); and with b12 and F105, recognition occurs with the less-protruding CDR loops and on one side of the extended CDR H3 (7, 48). Likewise, the presence of protruding CDR H3s observed in the quaternary-specific antibodies suggests that the recognized quaternary V2-V3 regions may be spread out over a relatively large surface or have critical sites of contact that are separated on the viral spike.

Affinity maturation and antibody breadth. Broad and potent neutralizing anti-HIV-1 antibodies like PG16 and VRC01 show extensive affinity maturation in their gene sequences, with 20 to 30% of their V genes altered by somatic hypermutation from their genomic precursors (31, 40, 43, 47). Moreover, we previously showed with the PG9 and PG16 antibodies that the degree of affinity maturation correlates with breadth and potency (31). Interestingly, the degree of affinity maturation observed with the strain-specific 2909 fits this pattern: the 2909 gene sequence reveals only a small number of somatic mutations, with 10 of 99 residues (10%) altered in the Vh gene and 6 out of 96 changed (6%) in the Vl gene (see Fig. S2 in the supplemental material). Whether appropriate affinity maturation could increase the degree of neutralization breadth with 2909 is unclear. What is clear is a strong correlation between affinity maturation and neutralization breadth for this class of antibodies. Recent data from longitudinal studies of HIV-1-infected individuals, which span the time from seroconversion to elicitation of quaternary-specific antibodies of breadth, show that the increase in breadth occurs gradually over several years (29). It will be interesting to correlate this slow increase in breadth with the degree of affinity maturation of the quaternary-specific antibodies.

Signature for and elicitation of quaternary-specific antibodies. The results suggest a potential signature for quaternary-specific antibodies with extensive neutralization breadth: CDR H3 of over 20 amino acids in length, affinity maturation in V-gene region of 20% or greater, prediction of tyrosine sulfation in CDR H3, and perhaps VH3-heavy chain gene usage, as VH3 is the gene used in 2909, PG9, and PG16. Whether such a signature could be used to identify antibodies directly from peripheral blood mononuclear cells of HIV-1-infected individuals is unclear, but an exciting potential application of the findings is made here.

More important, perhaps, is the use of the structural and functional insights gleaned from analyses reported here for the elicitation of additional quaternary-specific antibodies of high breadth and potency—the HIV-1 vaccine. First, since the derived antibodies target quaternary epitopes, it is important to immunize with true oligomeric replicas of the functional viral spike. Second, envelope-based immunogens should derive from strains for which the V2 and V3 regions recognized by quaternary-specific antibodies utilize sequences commonly seen in primary isolates (e.g., not ones with a rare lysine at position 160 in V2). Third, antibody breadth appears to be a function of affinity maturation, so factors that enhance affinity maturation should be utilized. Fourth, because the primary

component of the paratope appears to be the extended CDR H3, ways to increase the elicitation of antibodies with long CDR H3s should be tried, and immunization should be carried out with animals capable of producing CDR H3s of over 20 amino acids in length. For example, CDR H3 lengths of up to 22 residues have been found in rabbits (35). Fifth, because tyrosine sulfation appears to play an important role, it would be good to have test animals with sulfotransferases of human specificity.

Finally, we note that while the long CDR H3 associated with this class of antibodies may appear to be infrequently produced, this class of quaternary-specific antibodies does appear to be one of the more commonly elicited broadly neutralizing responses observed in sera from HIV-1-infected individuals (41). This suggests that processes which produce long CDR H3s are not so rare, a suggestion supported by the ready isolation of quaternary-specific rhesus antibodies with long CDR H3s. Thus, a focus on increasing the likelihood of other infrequent events—such as increasing the frequency of antibodies with affinity maturation levels of 20% or higher—may enhance the overall elicitation frequency of the desired quaternary-specific antibodies.

ACKNOWLEDGMENTS

We thank members of the Structural Biology Section and Structural Bioinformatics Core at the NIH Vaccine Research Center for comments on the manuscript, M. Gawinowicz of the Columbia University Protein Core Facility for N-terminal sequencing, and J. Stuckey for assistance with figures.

A.C., X.W., J.Z., J.R.M., and P.D.K. designed research; A.C., G.A.N., S.O., M.P., X.W., Y.Y., B.Z., and J.Z. performed research; M.K.G., S.P., J.E.R., L.S., and S.Z.-P. contributed essential reagents; A.C., X.W., J.Z., J.R.M., and P.D.K. analyzed the data; A.C., X.W., J.Z., and P.D.K. wrote the first draft of the paper, on which all authors commented.

Support for this work was provided by the International AIDS Vaccine Initiative (IAVI), by the Intramural Research Program of the NIH, by grants from the NIH (AI36085 and AI 27742), and by research funds from the Department of Veterans Affairs. Use of sector 22 (Southeast Region Collaborative Access team) at the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract no W-31-109-Eng-38.

ADDENDUM IN PROOF

The structures of human antibody 2909 and rhesus antibody 2.5B were recently reported by Kong et al. (X. Kong et al., *AIDS Res. Hum. Retroviruses* **26**:A56–A57, 2010). Their Fab 2909 crystals and structure are virtually identical to those reported here; their Fab 2.5B structure, meanwhile, confirms many of the features of extended CDR H3 proposed here, though at 1.9-Å resolution. Overall, the two studies are highly complementary. Kong et al. used structures of human and rhesus antibodies to define conserved structural features, while we integrated structural analyses with sequence analysis and chimeric antibodies to dissect similarities and differences between antibodies able to recognize a quaternary structure-specific site of HIV-I vulnerability.

REFERENCES

1. Adams, P. D., et al. 2004. Recent developments in the PHENIX software for automated crystallographic structure determination. *J. Synchrotron Radiat.* **11**:53–55.

2. Allcorn, L. C., and A. C. Martin. 2002. SACS—self-maintaining database of antibody crystal structure information. *Bioinformatics* **18**:175–181.
3. Barouch, D. H., et al. 2005. A human T-cell leukemia virus type 1 regulatory element enhances the immunogenicity of human immunodeficiency virus type 1 DNA vaccines in mice and nonhuman primates. *J. Virol.* **79**:8828–8834.
4. Burke, V., et al. 2009. Structural basis of the cross-reactivity of genetically related human anti-HIV-1 mAbs: implications for design of V3-based immunogens. *Structure* **17**:1538–1546.
5. Chang, W. C., et al. 2009. Incorporating support vector machine for identifying protein tyrosine sulfation sites. *J. Comput. Chem.* **30**:2526–2537.
6. Chen, C. H., L. Jin, C. Zhu, S. Holz-Smith, and T. J. Matthews. 2001. Induction and characterization of neutralizing antibodies against a human immunodeficiency virus type 1 primary isolate. *J. Virol.* **75**:6700–6704.
7. Chen, L., et al. 2009. Structural basis of immune evasion at the site of CD4 attachment on HIV-1 gp120. *Science* **326**:1123–1127.
8. Chen, V. B., et al. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**:12–21.
9. Cho, M. W., M. K. Lee, C. H. Chen, T. Matthews, and M. A. Martin. 2000. Identification of gp120 regions targeted by a highly potent neutralizing antiserum elicited in a chimpanzee inoculated with a primary human immunodeficiency virus type 1 isolate. *J. Virol.* **74**:9749–9754.
10. Collaborative Computational Project, No. 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**:760–763.
11. DeLano, W. L. 2002. The PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA.
12. Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**:2126–2132.
13. Etemad-Moghadam, B., et al. 1998. Characterization of simian-human immunodeficiency virus envelope glycoprotein epitopes recognized by neutralizing antibodies from infected monkeys. *J. Virol.* **72**:8437–8445.
14. Gorny, M. K., et al. 2005. Identification of a new quaternary neutralizing epitope on human immunodeficiency virus type 1 virus particles. *J. Virol.* **79**:5232–5237.
15. Holm, L., and C. Sander. 1995. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**:478–480.
16. Honnen, W. J., et al. 2007. Type-specific epitopes targeted by monoclonal antibodies with exceptionally potent neutralizing activities for selected strains of human immunodeficiency virus type 1 map to a common region of the V2 domain of gp120 and differ only at single positions from the clade B consensus sequence. *J. Virol.* **81**:1424–1432.
17. Huang, C. C., et al. 2005. Structure of a V3-containing HIV-1 gp120 core. *Science* **310**:1025–1028.
18. Jones, T. A., J. Y. Zou, S. W. Cowan, and M. Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**(Pt. 2):110–119.
19. Kabat, E. A., T. T. Wu, H. M. Perry, K. S. Gottesman, and C. Foeller. 1991. Sequences of proteins of immunological interest, 5th ed. U.S. Department of Health and Human Service, National Institutes of Health, Bethesda, MD.
20. Kimura, T., X. H. Wang, C. Williams, S. Zolla-Pazner, and M. K. Gorny. 2009. Human monoclonal antibody 2909 binds to pseudovirions expressing trimers but not monomeric HIV-1 envelope proteins. *Hum. Antibodies* **18**: 35–40.
21. Kraft, Z., et al. 2007. Macaques infected with a CCR5-tropic simian/human immunodeficiency virus (SHIV) develop broadly reactive anti-HIV neutralizing antibodies. *J. Virol.* **81**:6402–6411.
22. Kwong, P. D., et al. 1999. Probability analysis of variational crystallization and its application to gp120, the exterior envelope glycoprotein of type 1 human immunodeficiency virus (HIV-1). *J. Biol. Chem.* **274**:4115–4123.
23. Kwong, P. D., et al. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**:648–659.
24. Larkin, M. A., et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**:2947–2948.
25. Leitner, T. F., et al. 2006. HIV Sequence Compendium 2005. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.
26. McCoy, A. J., et al. 2007. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**:658–674.
27. Monigatti, F., E. Gasteiger, A. Bairoch, and E. Jung. 2002. The Sulfator: predicting tyrosine sulfation sites in protein sequences. *Bioinformatics* **18**: 769–770.
28. Moore, K. L. 2003. The biology and enzymology of protein tyrosine O-sulfation. *J. Biol. Chem.* **278**:24243–24246.
29. Morris, L., et al. 2010. How does neutralization breadth develop in HIV infection?, abstr. 013, p. 65. 2010 Keystone HIV Vaccine Symp., Banff, Canada.
30. Otwinowski, Z., and W. Minor. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**:307–326.
31. Pancera, M., et al. 2010. Crystal structure of PG16 and chimeric dissection with somatically related PG9: structure-function analysis of two quaternary-specific antibodies that effectively neutralize HIV-1. *J. Virol.* **84**:8098–8110.
32. Pejchal, J., L. M. Walker, D. R. Burton, and I. A. Wilson. 2010. Characterization of heterogeneity in PG9 tyrosine sulfation. *AIDS Res. Hum. Retrovir.* **26**:A155–A156.
33. Pejchal, R., et al. 2010. Structure and function of broadly reactive antibody PG16 reveal an H3 subdomain that mediates potent neutralization of HIV-1. *Proc. Natl. Acad. Sci. U. S. A.* **107**:11483–11488.
34. Petrey, D., et al. 2003. Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. *Proteins* **53**(Suppl. 6):430–435.
35. Popkov, M., et al. 2003. Rabbit immune repertoires as sources for therapeutic monoclonal antibodies: the impact of kappa allotype-correlated variation in cysteine content on antibody libraries selected by phage display. *J. Mol. Biol.* **325**:325–335.
36. Robinson, J. E., et al. 2010. Quaternary epitope specificities of anti-HIV-1 neutralizing antibodies generated in rhesus macaques infected by the simian/human immunodeficiency virus SHIVSF162P4. *J. Virol.* **84**:3443–3453.
37. Seaman, M. S., et al. 2010. Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. *J. Virol.* **84**: 1439–1452.
38. Smith, K., et al. 2009. Rapid generation of fully human monoclonal antibodies specific to a vaccinating antigen. *Nat. Protoc.* **4**:372–384.
39. Stanfield, R. L., M. K. Gorny, C. Williams, S. Zolla-Pazner, and I. A. Wilson. 2004. Structural rationale for the broad neutralization of HIV-1 by human monoclonal antibody 447-52D. *Structure* **12**:193–204.
40. Walker, L. M., et al. 2009. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326**:285–289.
41. Walker, L. M., et al. 2010. A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals. *PLoS Pathog.* **6**:e1001028.
42. Wang, G., and R. L. Dunbrack, Jr. 2003. PISCES: a protein sequence culling server. *Bioinformatics* **19**:1589–1591.
43. Wu, X., et al. 2010. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329**:856–861.
44. Wu, X., et al. 2009. Mechanism of human immunodeficiency virus type 1 resistance to monoclonal antibody B12 that effectively targets the site of CD4 attachment. *J. Virol.* **83**:10892–10907.
45. Zhang, Y., and J. Skolnick. 2005. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **33**:2302–2309.
46. Zhou, H., and Y. Zhou. 2002. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. *Protein Sci.* **11**:2714–2726.
47. Zhou, T., et al. 2010. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329**:811–817.
48. Zhou, T., et al. 2007. Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* **445**:732–737.
49. Zolla-Pazner, S., and T. Cardozo. 2010. Structure-function relationships of HIV-1 envelope sequence-variable regions refocus vaccine design. *Nat. Rev. Immunol.* **10**:527–535.