Sequence-Conserved and Antibody-Accessible Sites in the V1V2 Domain of HIV-1 gp120 Envelope Protein

Evgeny Shmelkov,[1] Arsen Grigoryan,[1] Chavdar Krachmarov,[1] Ruben Abagyan,[2] and Timothy Cardozo[1]

Abstract

The immune-correlates analysis of the RV144 trial suggested that epitopes targeted by protective antibodies (Abs) reside in the V1V2 domain of gp120. We mapped V1V2 positional sequence variation onto the conserved V1V2 structural fold and showed that while most of the solvent-accessible V1V2 amino acids vary between strains, there are two accessible molecular surface regions that are conserved and also naturally antigenic. These sites may contain epitopes targeted by broadly cross-reactive anti-V1V2 antibodies.

THE RESULTS OF THE IMMUNE-CORRELATE[1] and sieve[2] analyses of the RV144 trial demonstrated that a molecular correlate of risk for HIV infection was located in the V1V2 domain of HIV's gp120, suggesting that epitopes targeted by protective vaccine-induced antibodies (Abs) are located in this domain. Some protective anti-V1V2 Abs detected in the RV144 trial appeared to cross-react with multiple subtypes.[3] A number of V1V2-specific monoclonal Abs (mAbs) have been characterized to date.[4–8] As numerous anti-V1V2 mAbs were also shown to be broadly cross-reactive with multiple gp120 variants,[4,5,8–11] V1V2 is likely to possess conserved structural elements. Crystal structures of V1V2 from different viral strains obtained in complex with the broadly neutralizing Abs (BNAbs) PG9 and PG16 showed that this domain indeed forms a conserved four-stranded β-sheet fold in a Greek-key topology.[12,13] The crystallographic structure of the gp140 trimer affirmed that the same fold is also preserved in gp120-gp41 trimeric spikes from a third distinct strain.[14] Here, we combined these three-dimensional structural data with molecular epidemiological data from the LANL HIV database in order to understand how the strain-to-strain amino acid variability in this domain aligns with its structural preference in the broader context of the entire set of circulating HIV-1 strains.

In contrast to the V3 loop, V1V2 demonstrates substantial length variability. Analysis of the distribution of insertions and deletions in V1V2 (Supplementary Methods; Supplementary Data are available online at www.liebertpub.com/aid) suggests that V1V2 length polymorphism is clustered into two segments: the central positions in V1 (V1V2$^{132–153}$, mean length 23 amino acids) and the positions just C-terminal to the $\alpha_4\beta_7$ site in V2 (V1V2$^{187–188}$, mean length 6 amino acids; Fig. 1). Thus, it is likely that these segments are structurally polymorphic between strains. Moreover, the fact that one or both of these two segments (depending on the strain) were not resolved in the available V1V2 crystal structures is also indicative of their conformational flexibility. Both the length and conformational variation likely predispose these segments to immune escape, and antibodies targeting them are likely to be narrowly specific or type specific. Length variation at any other position in V1V2 is very rare, suggesting that immune escape there occurs primarily by varying side chain composition or by masking with nearby trimer peptide or glycan elements, as opposed to peptide backbone structural rearrangement.[15,16]

Accordingly, we calculated the side chain variability at each position of the V1V2 (Fig. 2; Supplementary Methods). The majority (29 out of 31) of V1V2 positions with variability scores higher than 50% are clustered to three linear segments: V1$^{133–152}$, V2$^{169–172}$, and V2$^{185–190}$, which are likely subject to immune pressure. Indeed, RV144-associated mAbs CH58/CH59[6] as well as the BNAbs PG9/PG16[8] target the V2$^{169–172}$ variable site ("C" ß-strand of V1V2). Narrowly cross-reactive CH58/CH59 mAbs engage variable amino acid side chains in this segment. In contrast, PG9/PG16 make sequence-independent contacts with the peptide backbone and a glycan, demonstrating, therefore, much broader cross-reactivity. CAP256 antibodies, while being broadly neutralizing for subtypes A and C, surprisingly target the highly variable lysine at position 169.[11,17] This could be explained by the fact that K^{169} is conserved in subtype C and K/R169 are conserved in subtype A (lysine and arginine have similar structural properties), but not in other subtypes.

[1]Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, New York.
[2]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California.
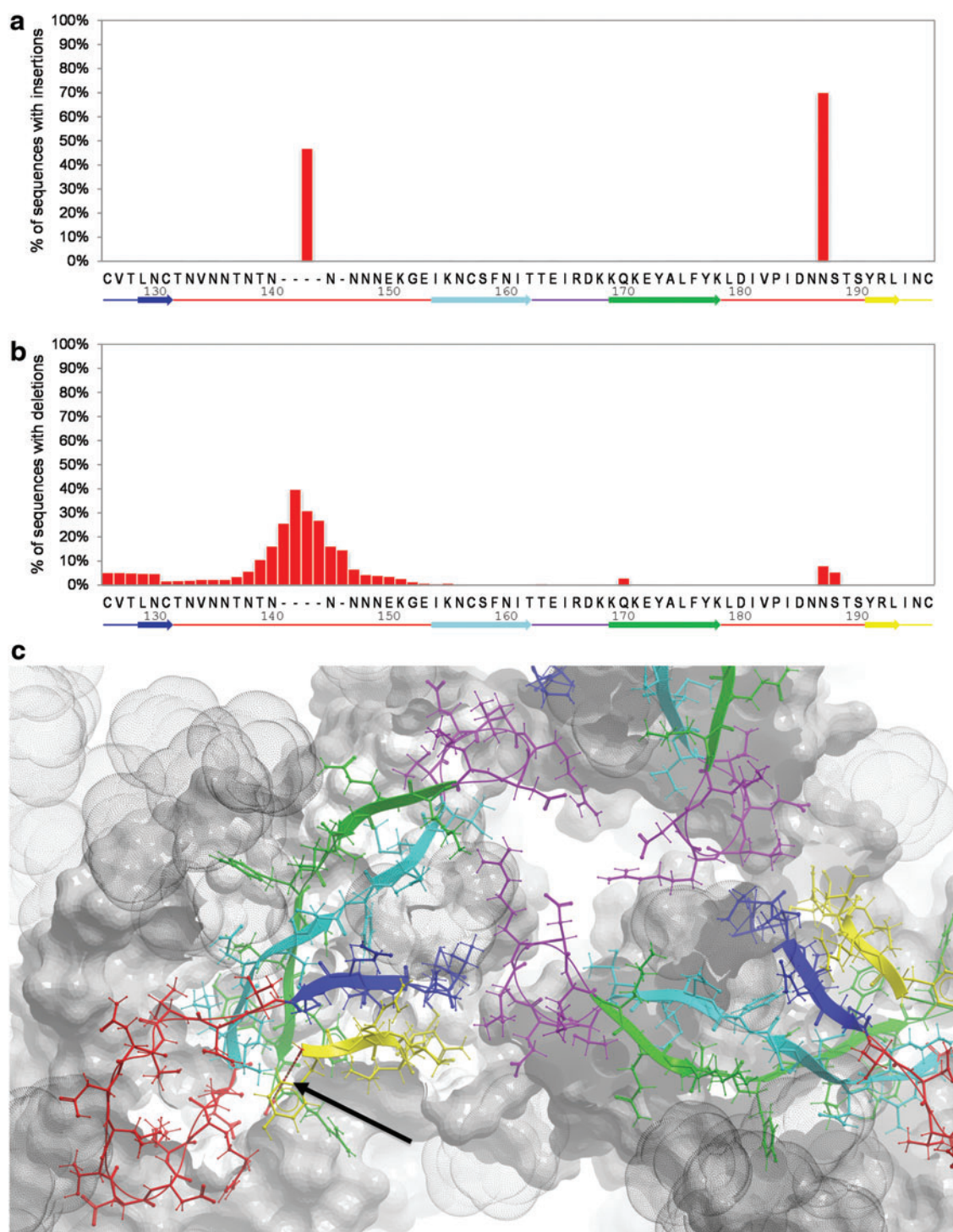
FIG. 1. Occurrence of insertions and deletions in the V1V2 domain. Distributions of insertions **(A)** and deletions **(B)** are shown for HIV-1 reference strain HxB2 positions from 126 to 196. The most frequently occurring side chains at each position corresponding to the HxB2 numbering are labeled on the *x*-axis. A linear secondary structure diagram of the V1V2 domain is provided at the *bottom* of each histogram. **(C)** Two clusters of high length polymorphism (colored in *red*) are mapped on the three-dimensional structure of the gp140 trimer. Note that the structure of the second cluster has not been resolved in the crystallographic study of the gp140 trimer and thus is shown as a *dashed red line* (*arrow*). The structure of V1V2 is shown in ribbon and stick representation (colors correspond to those on the secondary structure diagram in the *top* and *middle panels*), while the rest of the trimer is shown in skin representation. The space occupied by glycans is shown by *gray* dot-envelopes. Color images available online at www.liebertpub.com/aid
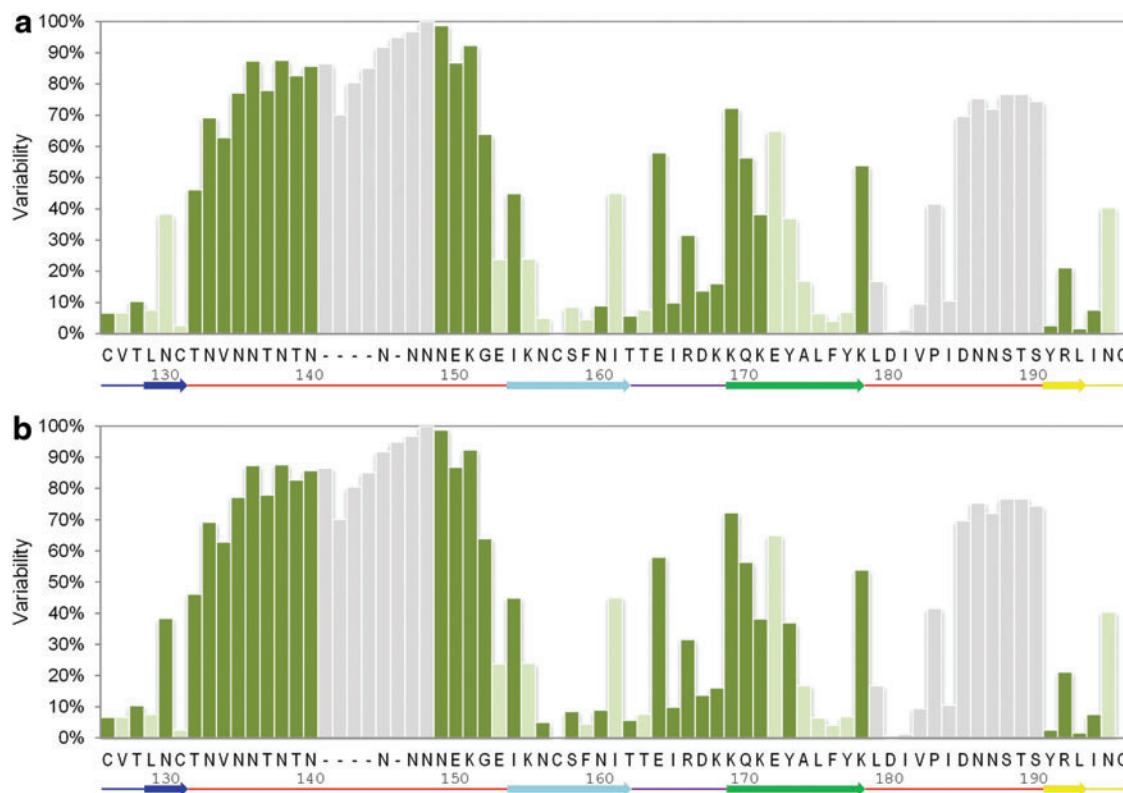
FIG. 2. Position-specific variability and accessibility in the V1V2 domain. HIV-1 reference strain HxB2 positions from 126 to 196 are shown. The most frequently occurring side chains at each position corresponding to the HxB2 numbering are labeled on the *x*-axis. Each bar is colored according to the mean relative accessibility of that specific side chain in the context of glycosylated **(a)** or nonglycosylated **(b)** gp140 trimer: *light green* for buried side chains (relative solvent accessibility <20%) and *dark green* for exposed side chains (relative solvent accessibility >20%). Those HxB2 positions that were deleted in the BG505 strain as well as those that were not resolved in the crystal structure are colored in *gray*. A linear secondary structure diagram of the V1V2 domain is provided at the bottom of the histogram. Color images available online at www.liebertpub.com/aid

Since these sequence variable regions are targeted by known Abs, they are expected to be solvent exposed. To test this, we calculated the solvent accessible area of each V1V2 amino acid in the context of the gp140 trimer (see Supplementary Methods). Since glycosylation could differ substantially between strains, we studied the accessibility of V1V2 both in a fully glycosylated trimer and in a trimer with removed glycans. As expected, the sequence-variable segments are exposed, while many of the sequence-conserved segments are poorly accessible (Fig. 2). The sequence-variable C-strand is the most accessible area of V1V2.

Interestingly, there are two relatively conserved and yet solvent-exposed V1V2 regions (Fig. 3). These regions are primarily composed of solvent-exposed amino acids (relative solvent accessibility of at least 20%; see Supplementary Methods) with variability scores of less than 50%, which are clustered together in a three-dimensional space. Region ''I'' is an intermonomer quaternary site formed by the tip of the V2 loop between its B and C strands (V2$^{162–168}$) and stems of V1V2 on the A and B strands (V1$^{127–128}$, V2$^{192–194}$) of an adjacent monomer. This region is located at the apex of the trimer relatively far (~ 15 Å) from the length-polymorphic V1V2 segments, and thus is likely accessible in many different strains. Region ''II'' is formed by positions located in three-dimensional proximity to the integrin binding site

(V2$^{176,179–184,191–194}$). This region can potentially be occluded by the two sequence-variable and structurally polymorphic loops in V1V2.

To test this, we performed computational sampling of one of these loops—V1$^{132–153}$ (see Supplementary Methods). Indeed, the conformational flexibility of V1 is high enough to partly mask epitopes in region ''II'' (Fig. 4). The length-polymorphic loop of V2 is expected to be at the same three-dimensional location (although not resolved in crystal structures), and likely contributes further to this masking. Interestingly, if all glycans are removed some additional conserved side chains become accessible (Supplementary Fig. S1). Some of them (V1V2^{130,156,158,160}) may further contribute to the formation of region ''I'' (Fig. 3). It is also important to mention that both regions ''I'' and ''II'' are conserved throughout subtypes A, B, C, and CRFs 01_AE and 02_AG (Supplementary Fig. S2; other clades were not studied due to a small sample size). This observation further emphasizes the potentially broadly cross-reactive nature of Abs targeting these two regions.

A linear fragment of region ''I'' (V2$^{165–168}$) has been previously confirmed as antigenic in a murine model.[7] CAP256 antibodies also depend on conserved positions at that same region.[11,17] On the other hand, seven human conformational anti-V2 mAb have been recently mapped to
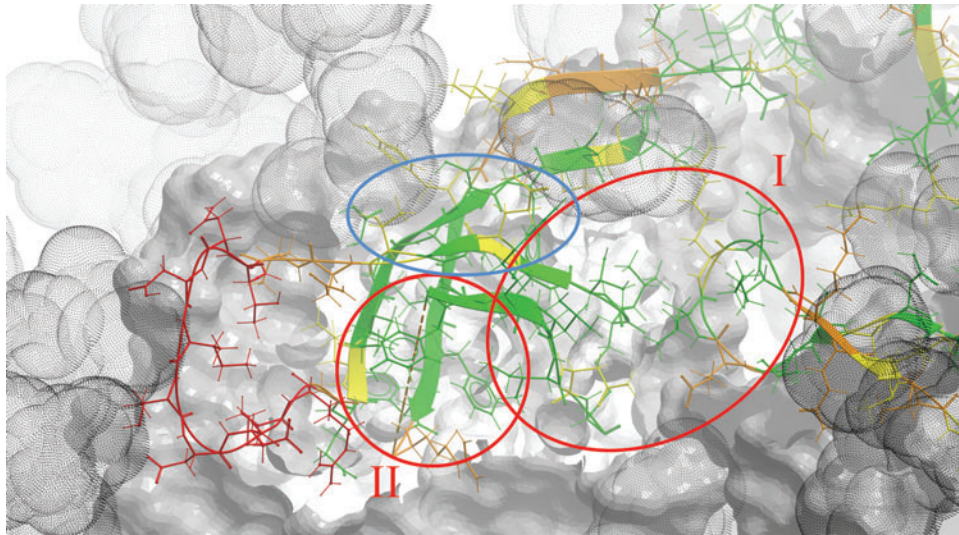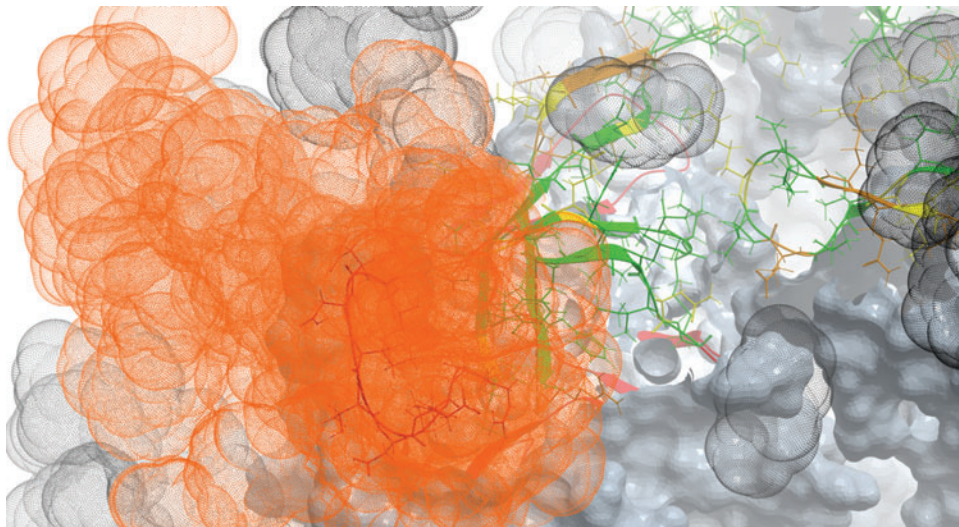
FIG. 3. Positional variability in V1V2 mapped onto the three-dimensional structure of the gp140 trimer. The structure of V1V2 (positions 126–196) is shown in ribbon and stick representation, while the rest of the trimer is shown in skin representation. Each V1V2 side chain is colored from *green* for sequence-conserved to *red* for sequence-variable positions. The space occupied by glycans is shown by *gray* dot-envelopes. The two epitope regions discussed in the text are highlighted with *red circles* and labeled as ''I'' and ''II.'' Conserved side chains that can contribute to the formation of region ''I'' in less glycosylated strains are highlighted with a *blue circle*. Color images available online at www.liebertpub.com/aid

V1V2 positions located within region ''II.''[18] Normally, it would be reasonable to expect high sequence diversity in exposed and naturally antigenic viral regions due to the constant immune pressure. However, both regions ''I'' and ''II'' are conserved. Therefore, these V1V2 regions could be sites of vulnerability in HIV.

If immune pressure so evidently leads to sequence and length variation in most of the Ab-exposed regions of V1V2, then why might regions ''I'' and ''II'' be conserved despite their exposure? The most obvious reason would be that al-

terations in these regions decrease viral fitness. Region ''I'' offers a ready explanation for this: sequence variation may perturb an intermonomer contact important for trimer integrity. Reasons for the conservation of region ''II'' are not obvious, but predicted partial masking of ''II'' by the length-variable loops may be one of them.

In summary, our map of sequence variability outlines two conserved, unmasked, and apparently antigenic V1V2 segments that could be sites of vulnerability to human host Abs. Region ''I'' is an especially promising site in being

FIG. 4. Conformational variability of the V1 loop in the context of the gp140 trimer. The structure of V1V2 (positions 126–196) is shown in ribbon and stick representation, while the rest of the trimer is shown in skin representation. Each V1V2 side chain is colored from *green* for sequence-conserved to *red* for sequence-variable positions. The space occupied by glycans is shown by *gray* dot-envelopes. The space that can reasonably be occupied by energetically favorable conformational flexibility of the crown segment of the V1 loop and the attached N137-linked glycan is shown as an *orange* cloud. Color images available online at www.liebertpub.com/aid

quaternary, unmasked, and sequence conserved. The map also shows that the C-strand of V1V2 is exposed, highly variable, and perhaps the most immunogenic region of V1V2. However, to be broadly cross-reactive, Abs targeting this region have to bind the structurally conserved peptide backbone atoms, some conserved side chains, and/or glycans.

If the protective anti-V1V2 Abs detected in the RV144 trial that cross-react with multiple subtypes[3] are clonal species, they may target one or both sites outlined here. Our delineation of the three-dimensional structures of these sites can be useful for identifying the Abs targeting these regions and for the rational design of antigens capable of eliciting broadly cross-reactive anti-V1V2 Abs.

Acknowledgments

Author Disclosure Statement

No competing financial interests exist.

References

1. Haynes BF, Gilbert PB, McElrath MJ, *et al.:* Immune-correlates analysis of an HIV-1 vaccine efficacy trial. N Engl J Med 2012;366(14):1275–1286.

2. Rolland M, Edlefsen PT, Larsen BB, *et al.:* Increased HIV-1 vaccine efficacy against viruses with genetic signatures in Env V2. Nature 2012;490(7420):417–420.

3. Zolla-Pazner S, deCamp A, Gilbert PB, *et al.:* Vaccine-induced IgG antibodies to V1V2 regions of multiple HIV-1 subtypes correlate with decreased risk of HIV-1 infection. PLoS One 2014;9(2):e87572.

4. Gorny MK, Moore JP, Conley AJ, *et al.:* Human anti-V2 monoclonal antibody that neutralizes primary but not laboratory isolates of human immunodeficiency virus type 1. J Virol 1994;68(12):8312–8320.

5. Gorny MK, Pan R, Williams C, *et al.:* Functional and immunochemical cross-reactivity of V2-specific monoclonal antibodies from HIV-1-infected individuals. Virology 2012;427(2):198–207.

6. Liao HX, Bonsignori M, Alam SM, *et al.:* Vaccine induction of antibodies against a structurally heterogeneous site of immune pressure within HIV-1 envelope protein variable regions 1 and 2. Immunity 2013;38(1):176–186.

7. Nakamura GR, Fonseca DP, O'Rourke SM, Vollrath AL, and Berman PW: Monoclonal antibodies to the V2 domain of MN-rgp120: Fine mapping of epitopes and inhibition of alpha4beta7 binding. PLoS One 2012;7(6):e39045.

8. Walker LM, Phogat SK, Chan-Hui PY, *et al.:* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science 2009;326(5950):285–289.

9. Shotton C, Arnold C, Sattentau Q, Sodroski J, and McKeating JA: Identification and characterization of monoclonal antibodies specific for polymorphic antigenic determinants within the V2 region of the human immunodeficiency virus type 1 envelope glycoprotein. J Virol 1995;69(1):222–230.

10. Corti D, Langedijk JP, Hinz A, *et al.:* Analysis of memory B cell responses and isolation of novel monoclonal antibodies with neutralizing breadth from HIV-1-infected individuals. PLoS One 2010;5(1):e8805.

11. Moore PL, Gray ES, Sheward D, *et al.:* Potent and broad neutralization of HIV-1 subtype C by plasma antibodies targeting a quaternary epitope including residues in the V2 loop. J Virol 2011;85(7):3128–3141.

12. McLellan JS, Pancera M, Carrico C, *et al.:* Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature 2011;480(7377):336–343.

13. Pancera M, Shahzad-Ul-Hussan S, Doria-Rose NA, *et al.:* Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. Nat Struct Mol Biol 2013;20(7):804–813.

14. Julien JP, Cupo A, Sok D, *et al.:* Crystal structure of a soluble cleaved HIV-1 envelope trimer. Science 2013;342:1477–1483.

15. Krachmarov C, Lai Z, Honnen WJ, *et al.:* Characterization of structural features and diversity of variable-region determinants of related quaternary epitopes recognized by human and rhesus macaque monoclonal antibodies possessing unusually potent neutralizing activities. J Virol 2011;85(20):10730–10740.

16. Krachmarov CP, Honnen WJ, Kayman SC, Gorny MK, Zolla-Pazner S, and Pinter A: Factors determining the breadth and potency of neutralization by V3-specific human monoclonal antibodies derived from subjects infected with clade A or clade B strains of human immunodeficiency virus type 1. J Virol 2006;80(14):7127–7135.

17. Moore PL, Sheward D, Nonyane M, *et al.:* Multiple pathways of escape from HIV broadly cross-neutralizing V2-dependent antibodies. J Virol 2013;87(9):4882–4894.

18. Mayr LM, Cohen S, Spurrier B, Kong XP, and Zolla-Pazner S: Epitope mapping of conformational V2-specific anti-HIV human monoclonal antibodies reveals an immunodominant site in V2. PLoS One 2013;8(7):e70859.

Address correspondence to:
Evgeny Shmelkov
180 Varick Street, Room 641
New York, New York 10014

E-mail: evs234@nyu.edu